

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Inouye, Sumiko  
Hsu, Mei-Yin  
Eagle, Susan  
Inouye, Masayori

(ii) TITLE OF INVENTION: Prokaryotic Reverse Transcriptase

(iii) NUMBER OF SEQUENCES: 45

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Weiser & Associates  
(B) STREET: 230 South Fifteenth Street, Suite 500  
(C) CITY: Philadelphia  
(D) STATE: Pennsylvania  
(E) COUNTRY: U.S.A.  
(F) ZIP: 19102

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/269,118  
(B) FILING DATE: 30-JUN-1994  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Weiser, Gerard J.  
(B) REGISTRATION NUMBER: 19,763  
(C) REFERENCE/DOCKET NUMBER: 377.5888P

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 215-875-8383  
(B) TELEFAX: 215-875-8394

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2176 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: CDS

26908037-00000000

(B) LOCATION: 640..2094

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCATCCGCGC GGACACCCCC TCCTACGTGC CCCCCGACGC GGAGAGCGGC GTGGAGACGG 60  
TGTACCGCGT TTCCCTGGAT GGTACACCTGG TGGCGGTGGA GTGGGGCCCCG CGCACGGGCT 120  
CGCCGCGTCA CCAGCGGCTC TGGTTCGACT CGGATGCGGA AGCCCCCGGA GCCTACTTCG 180  
CGCGCCTCGA GAAGTTGGCG GCTGACGGCT ACATCGACGC GGCCTCGGCA TTGGTCTAAA 240  
CCCTTCAACC ACGGCTCGGC CGCCACGCGC GGCCGGCAGG ACAGGTGCGA CGAACAGACG 300  
ACGACGTGCG CTTACGCGC GAGCAGCCGA GAGAGGTCCG GAGTGCATCA GCCTGAGCGC 360  
CTCGAGCGGC GGAGCGGCGT TCGCCGCTC CGGTTGGAAT GCAGGACACT CTCCGCAAGG 420  
TAGCCTGTTC TTGGCTCTCT CCCTCCTAGG CACTACGGCC AGGGTGGGTA GCGGAGCCAA 480  
CGACGCCACC GCCGTTTACC CACCCCGGCC GTAGTGCCTA GGAGGGGAGA GCCGGTGAGG 540  
CTACCGTGCC CCAGGTAAGA TGGTGGTGCT TTCCCGGCCT CCGTCGACTG CTCGCGCCAT 600  
GTCCCGTCTT CCATCGCCGC GCCCGCCCAA GGTGCAGAC ATG ACC GCC AGG CTG 654  
Met Thr Ala Arg Leu  
1 5  
GAC CCG TTC GTC CCC GCA GCT TCG CCG CAG GCC GTG CCC ACG CCC GAG 702  
Asp Pro Phe Val Pro Ala Ala Ser Pro Gln Ala Val Pro Thr Pro Glu  
10 15 20  
CTC ACC GCT CCG TCG TCA GAC GCG GCC GCG AAG CGT GAA GCC CGC CGG 750  
Leu Thr Ala Pro Ser Ser Asp Ala Ala Lys Arg Glu Ala Arg Arg  
25 30 35  
CTC GCG CAC GAA GCG TTG CTC GTC CGC GCG AAG GCC ATC GAC GAA GCG 798  
Leu Ala His Glu Ala Leu Leu Val Arg Ala Lys Ala Ile Asp Glu Ala  
40 45 50  
GGC GGC GCC GAC GAC TGG GTG CAG GCG CAG CTC GTC TCC AAG GGG CTC 846  
Gly Gly Ala Asp Asp Trp Val Gln Ala Gln Leu Val Ser Lys Gly Leu  
55 60 65  
GCG GTC GAG GAC CTG GAC TTC TCC AGC GCC TCC GAG AAG GAC AAG AAG 894  
Ala Val Glu Asp Leu Asp Phe Ser Ser Ala Ser Glu Lys Asp Lys Lys  
70 75 80 85  
GCC TGG AAG GAG AAG AAG AAG GCC GAG GCC ACC GAG CGC CGC GCG CTG 942  
Ala Trp Lys Glu Lys Lys Lys Ala Glu Ala Thr Glu Arg Arg Ala Leu  
90 95 100  
AAG CGT CAG GCG CAC GAG GCG TGG AAG GCC ACG CAC GTG GGC CAC CTG 990

Lys	Arg	Gln	Ala	His	Glu	Ala	Trp	Lys	Ala	Thr	His	Val	Gly	His	Leu	
			105					110					115			
GGC	GCG	GGC	GTG	CAC	TGG	GCG	GAG	GAC	CGC	CTG	GCC	GAC	GCG	TTC	GAC	1038
Gly	Ala	Gly	Val	His	Trp	Ala	Glu	Asp	Arg	Leu	Ala	Asp	Ala	Phe	Asp	
		120					125					130				
GTG	CCC	CAC	CGC	GAG	GAG	CGC	GCC	CGG	GCC	AAC	GGC	CTG	ACG	GAG	CTG	1086
Val	Pro	His	Arg	Glu	Glu	Arg	Ala	Arg	Ala	Asn	Gly	Leu	Thr	Glu	Leu	
	135					140					145					
GAC	TCC	GCG	GAG	GCG	CTG	GCC	AAG	GCG	CTG	GGG	CTG	AGC	GTC	TCC	AAG	1134
Asp	Ser	Ala	Glu	Ala	Leu	Ala	Lys	Ala	Leu	Gly	Leu	Ser	Val	Ser	Lys	
150					155					160					165	
CTC	CGC	TGG	TTC	GCG	TTC	CAC	CGG	GAG	GTC	GAC	ACG	GCC	ACG	CAC	TAC	1182
Leu	Arg	Trp	Phe	Ala	Phe	His	Arg	Glu	Val	Asp	Thr	Ala	Thr	His	Tyr	
			170					175						180		
GTG	AGC	TGG	ACC	ATT	CCG	AAG	CGG	GAC	GGC	AGC	AAG	CGC	ACG	ATT	ACG	1230
Val	Ser	Trp	Thr	Ile	Pro	Lys	Arg	Asp	Gly	Ser	Lys	Arg	Thr	Ile	Thr	
			185					190					195			
TCC	CCC	AAG	CCT	GAG	CTG	AAG	GCA	GCG	CAG	CGC	TGG	GTG	CTG	TCC	AAC	1278
Ser	Pro	Lys	Pro	Glu	Leu	Lys	Ala	Ala	Gln	Arg	Trp	Val	Leu	Ser	Asn	
		200					205					210				
GTC	GTG	GAG	CGG	CTG	CCG	GTC	CAC	GGC	GCC	GCC	CAC	GGC	TTC	GTG	GCG	1326
Val	Val	Glu	Arg	Leu	Pro	Val	His	Gly	Ala	Ala	His	Gly	Phe	Val	Ala	
	215					220					225					
GGA	CGC	TCC	ATC	CTC	ACC	AAC	GCG	CTG	GCC	CAC	CAG	GGC	GCG	GAC	GTC	1374
Gly	Arg	Ser	Ile	Leu	Thr	Asn	Ala	Leu	Ala	His	Gln	Gly	Ala	Asp	Val	
230					235					240					245	
GTG	GTC	AAG	GTG	GAC	CTC	AAG	GAC	TTC	TTC	CCC	TCC	GTC	ACC	TGG	CGC	1422
Val	Val	Lys	Val	Asp	Leu	Lys	Asp	Phe	Phe	Pro	Ser	Val	Thr	Trp	Arg	
			250					255						260		
CGG	GTG	AAG	GGC	CTG	TTG	CGC	AAG	GGC	GGC	CTG	CGG	CAG	GGC	ACG	TCC	1470
Arg	Val	Lys	Gly	Leu	Leu	Arg	Lys	Gly	Gly	Leu	Arg	Glu	Gly	Thr	Ser	
			265					270				275				
ACG	CTG	CTG	TCC	CTC	CTC	TCC	ACG	GAA	GCG	CCG	CGG	GAG	GCG	GTC	CAG	1518
Thr	Leu	Leu	Ser	Leu	Leu	Ser	Thr	Glu	Ala	Pro	Arg	Glu	Ala	Val	Gln	
		280					285					290				
TTC	CGC	GGC	AAG	CTC	CTG	CAC	GTC	GCC	AAG	GGC	CCG	CGC	GCC	CTG	CCC	1566
Phe	Arg	Gly	Lys	Leu	Leu	His	Val	Ala	Lys	Gly	Pro	Arg	Ala	Leu	Pro	
	295					300					305					
CAG	GGC	GCC	CCC	ACG	TCG	CCC	GGC	ATC	ACC	AAC	GCG	CTC	TGC	CTG	AAG	1614
Gln	Gly	Ala	Pro	Thr	Ser	Pro	Gly	Ile	Thr	Asn						

CTC	GAC	AAG	CGG	CTG	TCC	GCC	CTC	GCG	AAG	CGG	CTG	GGC	TTC	ACC	TAC	1662
Leu	Asp	Lys	Arg	Leu	Ser	Ala	Leu	Ala	Lys	Arg	Leu	Gly	Phe	Thr	Tyr	
			330						335					340		
ACG	CGC	TAC	GCG	GAC	GAC	CTG	ACC	TTC	TCC	TGG	ACG	AAG	GCG	AAG	CAG	1710
Thr	Arg	Tyr	Ala	Asp	Asp	Leu	Thr	Phe	Ser	Trp	Thr	Lys	Ala	Lys	Gln	
			345					350					355			
CCC	AAG	CCG	CGG	CGG	ACG	CAG	CGT	CCC	CCC	GTC	GCG	GTC	CTC	CTG	TCT	1758
Pro	Lys	Pro	Arg	Arg	Thr	Gln	Arg	Pro	Pro	Val	Ala	Val	Leu	Leu	Ser	
		360					365					370				
CGC	GTC	CAG	GAA	GTG	GTG	GAG	GCG	GAG	GGC	TTC	CGC	GTG	CAC	CCG	GAC	1806
Arg	Val	Gln	Glu	Val	Val	Glu	Ala	Glu	Gly	Phe	Arg	Val	His	Pro	Asp	
	375					380					385					
AAG	ACG	CGC	GTC	GCC	CGC	AAG	GGC	ACG	CGG	CAG	CGG	GTC	ACC	GGG	CTC	1854
Lys	Thr	Arg	Val	Ala	Arg	Lys	Gly	Thr	Arg	Gln	Arg	Val	Thr	Gly	Leu	
	390				395					400					405	
GTC	GTG	AAT	GCG	GCG	GGC	AAG	GAC	GCG	CCC	GCG	GCC	CGA	GTC	CCG	CGC	1902
Val	Val	Asn	Ala	Ala	Gly	Lys	Asp	Ala	Pro	Ala	Ala	Arg	Val	Pro	Arg	
			410					415						420		
GAC	GTC	GTC	CGC	CAG	CTC	CGC	GCC	GCC	ATC	CAC	AAC	CGG	AAG	AAG	GGC	1950
Asp	Val	Val	Arg	Gln	Leu	Arg	Ala	Ala	Ile	His	Asn	Arg	Lys	Lys	Gly	
			425				430						435			
AAG	CCG	GGC	CGC	GAG	GGC	GAG	TCG	CTC	GAG	CAG	CTC	AAG	GGC	ATG	GCC	1998
Lys	Pro	Gly	Arg	Glu	Gly	Glu	Ser	Leu	Glu	Gln	Leu	Lys	Gly	Met	Ala	
		440					445					450				
GCC	TTC	ATC	CAC	ATG	ACG	GAC	CCG	GCC	AAG	GGC	CGC	GCC	TTC	CTG	GCT	2046
Ala	Phe	Ile	His	Met	Thr	Asp	Pro	Ala	Lys	Gly	Arg	Ala	Phe	Leu	Ala	
	455					460					465					
CAG	CTC	ACG	GAG	CTC	GAG	TCC	ACG	GCG	AGC	GCC	GCT	CCG	CAG	GCG	GAG	2094
Gln	Leu	Thr	Glu	Leu	Glu	Ser	Thr	Ala	Ser	Ala	Ala	Pro	Gln	Ala	Glu	
	470				475				480						485	
TGACGCTCAG CGCGCGTCCG TCGCCGACGT GCCGCGCGCC AGCAACGCCG CATTCAGCAA																2154
CTCCGTCAGC CGGCGCGGGT AC																2176

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 263 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro  
1 5 10 15  
Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met  
20 25 30  
Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn  
35 40 45  
Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys  
50 55 60  
Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu  
65 70 75 80  
Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser  
85 90 95  
Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp  
100 105 110  
Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn  
115 120 125  
Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp  
130 135 140  
Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu  
145 150 155 160  
Pro Phe Lys Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp  
165 170 175  
Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys  
180 185 190  
Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro  
195 200 205  
Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu  
210 215 220  
Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys  
225 230 235 240  
Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn  
245 250 255  
Trp Ala Ser Gln Ile Tyr Pro  
260

09303031-000000

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 263 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Arg Pro Trp Ala Arg Thr Pro Pro Lys Ala Pro Arg Asn Gln Pro Val  
1 5 10 15  
Pro Phe Lys Pro Glu Arg Leu Gln Ala Leu Gln His Leu Val Arg Lys  
20 25 30  
Ala Leu Glu Ala Gly His Ile Glu Pro Tyr Thr Gly Pro Gly Asn Asn  
35 40 45  
Pro Val Phe Pro Val Lys Lys Ala Asn Gly Thr Trp Arg Phe Ile His  
50 55 60  
Asp Leu Arg Ala Thr Asn Ser Leu Thr Ile Asp Leu Ser Ser Ser Ser  
65 70 75 80  
Pro Gly Pro Pro Asp Leu Ser Ser Leu Pro Thr Thr Leu Ala His Leu  
85 90 95  
Gln Thr Ile Asp Leu Arg Asp Ala Phe Phe Gln Ile Pro Leu Pro Lys  
100 105 110  
Gln Phe Gln Pro Tyr Phe Ala Phe Thr Val Pro Gln Gln Cys Asn Tyr  
115 120 125  
Gly Pro Gly Thr Arg Tyr Ala Trp Lys Val Leu Pro Gln Gly Phe Lys  
130 135 140  
Asn Ser Pro Thr Leu Phe Glu Met Gln Leu Ala His Ile Leu Gln Pro  
145 150 155 160  
Ile Arg Gln Ala Phe Pro Gln Cys Thr Ile Leu Gln Tyr Met Asp Asp  
165 170 175  
Ile Leu Leu Ala Ser Pro Ser His Glu Asp Leu Leu Leu Leu Ser Glu  
180 185 190  
Ala Thr Met Ala Ser Leu Ile Ser His Gly Leu Pro Val Ser Glu Asn  
195 200 205  
Lys Thr Gln Gln Thr Pro Gly Thr Ile Lys Phe Leu Gly Gln Ile Ile  
210 215 220

DECEMBER 1, 1994  
Gent

Ser Pro Asn His Leu Thr Tyr Asp Ala Val Pro Thr Val Pro Ile Arg  
225 230 235 240

Ser Arg Trp Ala Leu Pro Glu Leu Gln Ala Leu Leu Gly Glu Ile Gln  
245 250 255

Trp Val Ser Lys Gly Thr Pro  
260

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 259 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Asn Val Leu Tyr Arg Ile Gly Ser Asp Asn Gln Tyr Thr Gln Phe Thr  
1 5 10 15

Ile Pro Lys Lys Gly Lys Gly Val Arg Thr Ile Ser Ala Pro Thr Asp  
20 25 30

Arg Leu Lys Asp Ile Gln Arg Arg Ile Cys Asp Leu Leu Ser Asp Cys  
35 40 45

Arg Asp Glu Ile Phe Ala Ile Arg Lys Ile Ser Asn Asn Tyr Ser Phe  
50 55 60

Gly Phe Glu Arg Gly Lys Ser Ile Ile Leu Asn Ala Tyr Lys His Arg  
65 70 75 80

Gly Lys Gln Ile Ile Leu Asn Ile Asp Leu Lys Asp Phe Phe Glu Ser  
85 90 95

Phe Asn Phe Gly Arg Val Arg Gly Tyr Phe Leu Ser Asn Gln Asp Phe  
100 105 110

Leu Leu Asn Pro Val Val Ala Thr Thr Leu Ala Lys Ala Ala Cys Tyr  
115 120 125

Asn Gly Thr Leu Pro Gln Gly Ser Pro Cys Ser Pro Ile Ile Ser Asn  
130 135 140

Leu Ile Cys Asn Ile Met Asp Met Arg Leu Ala Lys Leu Ala Lys Lys  
145 150 155 160

Tyr Gly Cys Thr Tyr Ser Arg Tyr Ala Asp Asp Ile Thr Ile Ser Thr  
165 170 175

030303031.030303030  
Sue  
G  
cont

Asn Lys Asn Thr Phe Pro Leu Glu Met Ala Thr Val Gln Pro Glu Gly  
 180 185 190  
 Val Val Leu Gly Lys Val Leu Val Lys Glu Ile Glu Asn Ser Gly Phe  
 195 200 205  
 Glu Ile Asn Asp Ser Lys Thr Arg Leu Thr Tyr Lys Thr Ser Arg Gln  
 210 215 220  
 Glu Val Thr Gly Leu Thr Val Asn Arg Ile Val Asn Ile Asp Arg Cys  
 225 230 235 240  
 Tyr Tyr Lys Lys Thr Arg Ala Leu Ala His Ala Leu Tyr Arg Thr Gly  
 245 250 255  
 Glu Tyr Lys

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 266 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ala Phe His Arg Glu Val Asp Thr Ala Thr His Tyr Val Ser Trp Thr  
 1 5 10 15  
 Ile Pro Lys Arg Asp Gly Ser Lys Arg Thr Ile Thr Ser Pro Lys Pro  
 20 25 30  
 Glu Leu Lys Ala Ala Gln Arg Trp Val Leu Ser Asn Val Val Glu Arg  
 35 40 45  
 Leu Pro Val His Gly Ala Ala His Gly Phe Val Ala Gly Arg Ser Ile  
 50 55 60  
 Leu Thr Asn Ala Leu Ala His Gln Gly Ala Asp Val Val Val Lys Val  
 65 70 75 80  
 Asp Leu Lys Asp Phe Phe Pro Ser Val Thr Trp Arg Arg Val Lys Gly  
 85 90 95  
 Leu Leu Arg Lys Gly Gly Leu Arg Glu Gly Thr Ser Thr Leu Leu Ser  
 100 105 110  
 Leu Leu Ser Thr Glu Ala Pro Arg Glu Ala Val Gln Phe Arg Gly Lys  
 115 120 125

CY  
 G  
 cont.



Leu Leu His Val Ala Lys Gly Pro Arg Ala Leu Pro Gln Gly Ala Pro  
 130 135 140  
 Thr Ser Pro Gly Ile Thr Asn Ala Leu Cys Leu Lys Leu Asp Lys Arg  
 145 150 155 160  
 Leu Ser Ala Leu Ala Lys Arg Leu Gly Phe Thr Tyr Thr Arg Tyr Ala  
 165 170 175  
 Asp Asp Leu Thr Phe Ser Trp Thr Lys Ala Lys Gln Pro Lys Pro Arg  
 180 185 190  
 Arg Thr Gln Arg Pro Pro Val Ala Val Leu Leu Ser Arg Val Gln Glu  
 195 200 205  
 Val Val Glu Ala Glu Gly Phe Arg Val His Pro Asp Lys Thr Arg Val  
 210 215 220  
 Ala Arg Lys Gly Thr Arg Gln Arg Val Thr Gly Leu Val Val Asn Ala  
 225 230 235 240  
 Ala Gly Lys Asp Ala Pro Ala Ala Arg Val Pro Arg Asp Val Val Arg  
 245 250 255  
 Gln Leu Arg Ala Ala Ile His Asn Arg Lys  
 260 265

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 111 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Pro Thr Pro Glu Leu Thr Ala Pro Ser Ser Asp Ala Ala Ala Lys Arg  
 1 5 10 15  
 Glu Ala Arg Arg Leu Ala His Glu Ala Leu Leu Val Arg Ala Lys Ala  
 20 25 30  
 Ile Asp Glu Ala Gly Gly Ala Asp Asp Trp Val Gln Ala Gln Leu Val  
 35 40 45  
 Ser Lys Gly Leu Ala Val Glu Asp Leu Asp Phe Ser Ser Ala Ser Glu  
 50 55 60  
 Lys Asp Lys Lys Ala Trp Lys Glu Lys Lys Lys Ala Glu Ala Thr Glu  
 65 70 75 80

DELETED - 030000

See  
 G  
 cont

Arg Arg Ala Leu Lys Arg Gln Ala His Glu Ala Trp Lys Ala Thr His  
85 90 95

Val Gly His Leu Gly Ala Gly Val His Trp Ala Glu Asp Arg Leu  
100 105 110

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 110 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Pro Asp Pro Asp Met Thr Arg Val Thr Asn Ser Pro Ser Leu Gln Ala  
1 5 10 15  
His Leu Gln Ala Leu Tyr Leu Val Gln His Glu Val Trp Arg Pro Leu  
20 25 30  
Ala Ala Ala Tyr Gln Glu Gln Leu Asp Arg Pro Val Val Pro His Pro  
35 40 45  
Tyr Arg Val Gly Asp Thr Val Trp Val Arg Arg His Gln Thr Lys Asn  
50 55 60  
Leu Glu Pro Arg Trp Lys Gly Pro Tyr Thr Val Leu Leu Thr Thr Pro  
65 70 75 80  
Thr Ala Leu Lys Val Asp Gly Ile Ala Ala Trp Ile His Ala Ala His  
85 90 95  
Val Lys Ala Ala Asp Pro Gly Gly Gly Pro Ser Ser Arg Leu  
100 105 110

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Gly Lys Asp Ala Pro Ala Ala Arg Val Pro Arg Asp Val Val Arg Gln

RECEIVED - FEBRUARY 1981

Sub  
G  
cont

1				5					10					15		
Leu	Arg	Ala	Ala	Ile	His	Asn	Arg	Lys	Lys	Gly	Lys	Pro	Gly	Arg	Glu	
			20					25					30			
Gly	Glu	Ser	Leu	Glu	Gln	Leu	Lys	Gly	Met	Ala	Ala	Phe	Ile	His	Met	
		35					40					45				
Thr	Asp	Pro	Ala	Lys	Gly	Arg	Ala	Phe	Leu	Ala	Gln	Leu	Thr	Glu	Leu	
	50					55					60					
Glu	Ser	Thr	Ala	Ser	Ala	Ala	Pro	Gln	Ala	Glu						
65					70					75						

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Gly Lys Glu Gly His Ser Ala Arg Gln Cys Arg Ala Pro Arg Arg Gln  
1 5 10 15

Gly Cys Trp Lys Cys Gly Lys / Pro Gly His Ile Met Thr Asn Cys Pro  
20 25 30

Asp Arg Gln Ala Gly Phe Leu ~~Gly~~ Leu Gly Pro Trp Gly Lys Lys Pro  
35 40 45

Arg Asn Phe Pro Val Ala Gln Val Pro Gln Gly Leu Thr Pro Thr Ala  
50 55 60

Pro Pro  
65

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 68 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Gly Pro Arg Ala Leu Pro Gln Gly Ala Pro Thr Ser Pro Gly Ile Thr  
 1 5 10 15  
 Asn Ala Leu Cys Leu Lys Leu Asp Lys Arg Leu Ser Ala Leu Ala Lys  
 20 25 30  
 Arg Leu Gly Phe Thr Tyr Thr Arg Tyr Ala Asp Asp Leu Thr Phe Ser  
 35 40 45  
 Trp Thr Lys Ala Lys Gln Pro Lys Pro Arg Arg Thr Gln Arg Pro Pro  
 50 55 60  
 Val Ala Val Leu  
 65

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 68 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Tyr Asn Gly Thr Leu Pro Gln Gly Ser Pro Cys Ser Pro Ile Ile Ser  
 1 5 10 15  
 Asn Leu Ile Cys Asn Ile Met Asp Met Arg Leu Ala Lys Leu Ala Lys  
 20 25 30  
 Lys Tyr Gly Cys Thr Tyr Ser Arg Tyr Ala Asp Asp Ile Thr Ile Ser  
 35 40 45  
 Thr Asn Lys Asn Thr Phe Pro Leu Glu Met Ala Thr Val Gln Pro Glu  
 50 55 60  
 Gly Val Val Leu  
 65

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 68 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

RECORDED - 030382

Sub  
 G-1  
 Cont

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Tyr Lys Asn Leu Leu Pro Gln Gly Ala Pro Ser Ser Pro Lys Leu Ala  
1            5            10            15

Asn Leu Ile Cys Ser Lys Leu Asp Tyr Arg Ile Gln Gly Tyr Ala Gly  
          20            25            30

Ser Arg Gly Leu Ile Tyr Thr Arg Tyr Ala Asp Asp Leu Thr Leu Ser  
          35            40            45

Ala Gln Ser Met Lys Lys Val Val Lys Ala Arg Asp Phe Leu Phe Ser  
      50            55            60

Ile Ile Pro Ser  
65

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 67 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION:\ SEQ ID NO:13:

[illegible]

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 66 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

**SECRET**

(2) INFORMATION FOR SEQ ID NO:15:

(ii) MOLECULE TYPE: protein

[illegible]

(2) INFORMATION FOR SEQ ID NO:16:

(ii) MOLECULE TYPE: protein

~~000000~~ - 000007

[illegible]

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 65 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

[illegible]

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Ile	Ala	Thr	Asn	Gly	Val	Pro	Gln	Gly	Ala	Ser	Thr	Ser	Cys	Gly	Leu
1				5				10						15	
Ala	Thr	Tyr	Asn	Val	Leu	Glu	Leu	Phe	Leu	Arg	Tyr	Asp	Glu	Leu	Ile
			20					25					30		
Met	Tyr	Ala	Asp	Asp	Gly	Ile	Leu	Cys	Arg	Gln	Asp	Pro	Ser	Thr	Pro
		35					40					45			
Asp	Phe	Ser	Val	Glu	Glu	Ala	Gly	Val	Val	Gln	Glu	Pro			
	50					55					60				

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 61 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Tyr	Glu	Tyr	Leu	Arg	Met	Pro	Phe	Gly	Leu	Lys	Asn	Ala	Pro	Ala	Thr
1				5				10						15	
Phe	Gln	Arg	Cys	Met	Asn	Asp	Ile	Leu	Arg	Pro	Leu	Leu	Asn	Lys	His
			20					25					30		
Cys	Leu	Val	Tyr	Leu	Asp	Asp	Ile	Ile	Val	Phe	Ser	Thr	Ser	Leu	Asp
		35					40					45			
Glu	His	Leu	Gln	Ser	Leu	Gly	Leu	Val	Phe	Glu	Lys	Leu			
	50					55					60				

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 61 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

030001-030007  
260000-260000



[illegible]

069808 - 030357

06980007L-030357

068000 - 030357

- [illegible]

080609Z - 030357

069080. - 030357

0360801 - 030357

069080. - 030357

060908 - 030357Z

- [illegible]

060908Z - 030357

[illegible][illegible]

Trp Asn Ala Cys Val Tyr Val Lys Gln Val Ser Glu Gln Glu His Leu  
20 25 30  
Tyr Leu Leu Leu Tyr Val Asp Asp Met Leu Ile Ala Gly Lys Ser Lys  
35 40 45  
Ser Glu Ile Asn Lys Val Lys Glu Gln Leu Ser Met Glu Phe  
50 55 60

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 63 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Ile Arg Leu Lys Lys Ser Leu Tyr Glu Leu Lys Gln Ser Gly Ala Asn  
1 5 10 15  
Trp Tyr Glu Glu Val Arg Gly Trp Ser Cys Val Phe Lys Asn Ser Gln  
20 25 30  
Val Thr Ile Cys Leu Phe Val Asp Asp Met Val Leu Phe Ser Lys Asn  
35 40 45  
Leu Asn Ser Asn Lys Arg Ile Ile Glu Lys Leu Lys Met Gln Tyr  
50 55 60

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 58 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION: 15  
(D) OTHER INFORMATION: /note= "The 2' position of this  
nucleotide is linked to the 5' position of  
nucleotide number 1 of SEQ ID NO: 25 of this  
application."

(ix) FEATURE:

- (A) NAME/KEY: misc\_binding  
(B) LOCATION: 52..58

26E0E0"TE0E0E0E0

030309 J. 030309

CACGCAUGUA GGCAGAUUUG UUGGUUGUGA AUCGCAACCA GUGGCCUUA A UGGCAGGA

(2) INFORMATION FOR SEQ ID NO:25:

(ix) FEATURE:

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TCCTTCGCAC AGCACACCTG CCGTATAGCT CTGAATCAAG GATTTTAGGG AGGCGATTCC

60

TCCTGCC

67

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2423 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 418..2175

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TGGCCATTNA GATACGGATT TTCACTTCCT TGACAGTGCA TGACTATGCT GCATGAAATN 60  
GCATGATCGA TTGAGGATCG TCTTTGCTCA GATCCGCCAG AACTGGCGGG CTTTTGCTCA 120  
TGTCATGCAT GTGCATGAAA ACCACTGCAT AAAGCGGGCA GCGTGGCGG GGATACGAGC 180  
GCGCGCTATC ACCGAAAATA GCCAAAATAC TTCTGGAAAA CAGAAAGTTG AAGTGATATG 240  
TTCATAAACA CGCATGTAGG CAGATTTGTT GGTGTGAAT CGCAACCAGT GGCCTTAATG 300  
GCAGGAGGAA TCGCCTCCCT AAAATCCTTG ATTCAGAGCT ATACGGCAGG TGTGCTGTGC 360  
GAAGGAGTGC CTGCATGCGT TTCTCCTTGG CCTTTTTTCC TCTGGGATGA AGAAGAA 417  
ATG ACA AAA ACA TCT AAA CTT GAC GCA CTT AGG GCT GCT ACT TCA CGT 465  
Met Thr Lys Thr Ser Lys Leu Asp Ala Leu Arg Ala Ala Thr Ser Arg  
1 5 10 15  
GAA GAC TTG GCT AAA ATT TTA GAT ATT AAG TTG GTA TTT TTA ACT AAC 513  
Glu Asp Leu Ala Lys Ile Leu Asp Ile Lys Leu Val Phe Leu Thr Asn  
20 25 30  
GTT CTA TAT AGA ATC GGC TCG GAT AAT CAA TAC ACT CAA TTT ACA ATA 561  
Val Leu Tyr Arg Ile Gly Ser Asp Asn Gln Tyr Thr Gln Phe Thr Ile  
35 40 45  
CCG AAG AAA GGA AAA GGG GTA AGG ACT ATT TCT GCA CCT ACA GAC CGG 609  
Pro Lys Lys Gly Lys Gly Val Arg Thr Ile Ser Ala Pro Thr Asp Arg  
50 55 60  
TTG AAG GAC ATC CAA CGA AGA ATA TGT GAC TTA CTT TCT GAT TGT AGA 657  
Leu Lys Asp Ile Gln Arg Arg Ile Cys Asp Leu Leu Ser Asp Cys Arg  
65 70 75 80  
GAT GAG ATC TTT GCT ATA AGG AAA ATT AGT AAC AAC TAT TCC TTT GGT 705  
Asp Glu Ile Phe Ala Ile Arg Lys Ile Ser Asn Asn Tyr Ser Phe Gly  
85 90 95  
TTT GAG AGG GGA AAA TCA ATA ATC CTA AAT GCT TAT AAG CAT AGA GGC 753  
Phe Glu Arg Gly Lys Ser Ile Ile Leu Asn Ala Tyr Lys His Arg Gly  
100 105 110  
AAA CAA ATA ATA TTA AAT ATA GAT CTT AAG GAT TTT TTT GAA AGC TTT 801  
Lys Gln Ile Ile Leu Asn Ile Asp Leu Lys Asp Phe Phe Glu Ser Phe  
115 120 125  
AAT TTT GGA CGA GTT AGA GGA TAT TTT CTT TCC AAT CAG GAT TTT TTA 849  
Asn Phe Gly Arg Val Arg Gly Tyr Phe Leu Ser Asn Gln Asp Phe Leu  
130 135 140  
TTA AAT CCT GTG GTG GCA ACG ACA CTT GCA AAA GCT GCA TGC TAT AAT 897  
Leu Asn Pro Val Val Ala Thr Thr Leu Ala Lys Ala Ala Cys Tyr Asn

145					150					155					160			
GGA Gly	ACC Thr	CTC Leu	CCC Pro	CAA Gln 165	GGA Gly	AGT Ser	CCA Pro	TGT Cys	TCT Ser 170	CCT Pro	ATT Ile	ATC Ile	TCA Ser	AAT Asn 175	CTA Leu		945	
ATT Ile	TGC Cys	AAT Asn	ATT Ile 180	ATG Met	GAT Asp	ATG Met	AGA Arg	TTA Leu 185	GCT Ala	AAG Lys	CTG Leu	GCT Ala	AAA Lys 190	AAA Lys	TAT Tyr		993	
GGA Gly	TGT Cys	ACT Thr 195	TAT Tyr	AGC Ser	AGA Arg	TAT Tyr	GCT Ala 200	GAT Asp	GAT Asp	ATA Ile	ACA Thr	ATT Ile 205	TCT Ser	ACA Thr	AAT Asn		1041	
AAA Lys	AAT Asn 210	ACA Thr	TTT Phe	CCG Pro	TTA Leu	GAA Glu 215	ATG Met	GCT Ala	ACT Thr	GTG Val	CAA Gln 220	CCT Pro	GAA Glu	GGG Gly	GTT Val		1089	
GTT Val 225	TTG Leu	GGA Gly	AAA Lys	GTT Val	TTG Leu 230	GTA Val	AAA Lys	GAA Glu	ATA Ile	GAA Glu 235	AAC Asn	TCT Ser	GGA Gly	TTC Phe	GAA Glu 240		1137	
ATA Ile	AAT Asn	GAT Asp	TCA Ser	AAG Lys 245	ACT Thr	AGG Arg	CTT Leu	ACG Thr	TAT Tyr 250	AAG Lys	ACA Thr	TCA Ser	AGG Arg	CAA Gln 255	GAA Glu		1185	
GTA Val	ACG Thr	GGA Gly	CTT Leu 260	ACA Thr	GTT Val	AAC Asn	AGA Arg	ATC Ile 265	GTT Val	AAT Asn	ATT Ile	GAT Asp	AGA Arg 270	TGT Cys	TAT Tyr		1233	
TAT Tyr	AAA Lys	AAA Lys 275	ACT Thr	CGG Arg	GCG Ala	TTG Leu	GCA Ala 280	CAT His	GCT Ala	TTG Leu	TAT Tyr	CGT Arg 285	ACA Thr	GGT Gly	GAA Glu		1281	
TAT Tyr	AAA Lys 290	GTG Val	CCA Pro	GAT Asp	GAA Glu	AAT Asn 295	GGT Gly	GTT Val	TTA Leu	GTT Val	TCA Ser 300	GGA Gly	GGT Gly	CTG Leu	GAT Asp		1329	
AAA Lys 305	CTT Leu	GAG Glu	GGG Gly	ATG Met	TTT Phe 310	GGT Gly	TTT Phe	ATT Ile	GAT Asp	CAA Gln 315	GTT Val	GAT Asp	AAG Lys	TTT Phe	AAC Asn 320		1377	
AAT Asn	ATA Ile	AAG Lys	AAA Lys	AAA Lys 325	CTG Leu	AAC Asn	AAG Lys	CAA Gln	CCT Pro 330	GAT Asp	AGA Arg	TAT Tyr	GTA Val	TTG Leu 335	ACT Thr		1425	
AAT Asn	GCG Ala	ACT Thr	TTG Leu 340	CAT His	GGT Gly	TTT Phe	AAA Lys	TTA Leu 345	AAG Lys	TTG Leu	AAT Asn	GCG Ala	CGA Arg 350	GAA Glu	AAA Lys		1473	
GCA Ala	TAT Tyr	AGT Ser 355	AAA Lys	TTT Phe	ATT Ile	TAC Tyr	TAT Tyr 360	AAA Lys	TTT Phe	TTT Phe	CAT His	GGC Gly 365	AAC Asn	ACC Thr	TGT Cys		1521	

CCT Pro	ACG Thr 370	ATA Ile	ATT Ile	ACA Thr	GAA Glu	GGG Gly 375	AAG Lys	ACT Thr	GAT Asp	CGG Arg	ATA Ile 380	TAT Tyr	TTG Leu	AAG Lys	GCT Ala	1569
GCT Ala 385	TTG Leu	CAT His	TCT Ser	TTG Leu	GAG Glu 390	ACA Thr	TCA Ser	TAT Tyr	CCT Pro	GAG Glu 395	TTG Leu	TTT Phe	AGA Arg	GAA Glu	AAA Lys 400	1617
ACA Thr	GAT Asp	AGT Ser	AAA Lys	AAG Lys 405	AAA Lys	GAA Glu	ATA Ile	AAT Asn	CTT Leu 410	AAT Asn	ATA Ile	TTT Phe	AAA Lys	TCT Ser 415	AAT Asn	1665
GAA Glu	AAG Lys	ACC Thr	AAA Lys 420	TAT Tyr	TTT Phe	TTA Leu	GAT Asp	CTT Leu 425	TCT Ser	GGG Gly	GGA Gly	ACT Thr	GCA Ala 430	GAT Asp	CTG Leu	1713
AAA Lys	AAA Lys	TTT Phe 435	GTA Val	GAG Glu	CGT Arg	TAT Tyr	AAA Lys 440	AAT Asn	AAT Asn	TAT Tyr	GCT Ala	TCT Ser 445	TAT Tyr	TAT Tyr	GGT Gly	1761
TCT Ser	GTT Val 450	CCA Pro	AAA Lys	CAG Gln	CCA Pro	GTG Val 455	ATT Ile	ATG Met	GTT Val	CTT Leu	GAT Asp 460	AAT Asn	GAT Asp	ACA Thr	GGT Gly	1809
CCA Pro 465	AGC Ser	GAT Asp	TTA Leu	CTT Leu	AAT Asn 470	TTT Phe	CTG Leu	CGC Arg	AAT Asn	AAA Lys 475	GTT Val	AAA Lys	AGC Ser	TGC Cys	CCA Pro 480	1857
GAC Asp	GAT Asp	GTA Val	ACT Thr	GAA Glu 485	ATG Met	AGA Arg	AAG Lys	ATG Met	AAA Lys 490	TAT Tyr	ATT Ile	CAT His	GTT Val	TTC Phe 495	TAT Tyr	1905
AAT Asn	TTA Leu	TAT Tyr	ATA Ile 500	GTT Val	CTC Leu	ACA Thr	CCA Pro	TTG Leu 505	AGT Ser	CCT Pro	TCC Ser	GGC Gly	GAA Glu 510	CAA Gln	ACT Thr	1953
TCA Ser	ATG Met	GAG Glu 515	GAT Asp	CTT Leu	TTC Phe	CCT Pro	AAA Lys 520	GAT Asp	ATT Ile	TTA Leu	GAT Asp	ATC Ile 525	AAG Lys	ATT Ile	GAT Asp	2001
GGT Gly	AAG Lys 530	AAA Lys	TTC Phe	AAC Asn	AAA Lys	AAT Asn 535	AAT Asn	GAT Asp	GGA Gly	GAC Asp	TCA Ser 540	AAA Lys	ACG Thr	GAA Glu	TAT Tyr	2049
GGG Gly 545	AAG Lys	CAT His	ATT Ile	TTT Phe	TCC Ser 550	ATG Met	AGG Arg	GTT Val	GTT Val	AGA Arg 555	GAT Asp	AAA Lys	AAG Lys	CGG Arg	AAA Lys 560	2097
ATA Ile	GAT Asp	TTT Phe	AAG Lys	GCA Ala 565	TTT Phe	TGT Cys	TGT Cys	ATT Ile	TTT Phe 570	GAT Asp	GCT Ala	ATA Ile	AAA Lys	GAT Asp 575	ATA Ile	2145
AAG Lys	GAA Glu	CAT His	TAT Tyr	AAA Lys	TTA Leu	ATG Met	TTA Leu	AAT Asn	AGC Ser	TAATGAACAG CCCTAACGTT						2195

585

ATGAACGCTA	AGGCTGATT	TTCGTTAAAA	TTTATATGGT	TTGAATTGTA	ATATATTATC	2255
TTCAAGCCAT	TTATTTAATT	CCTGCATCCT	TTTCTGTAAG	GGTATTAATT	CGTTCCTCAC	2315
AAACACTAAA	CTCGCTTTTT	CCACATCCCC	AAACCCCCCT	AACATTATTC	GGCATAATCC	2375
CCATCATTTG	CGGTGGCACA	CGATGCGCTG	CCATCATGTC	ATCGCGGC		2423

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 546 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Val 1	Lys	Leu	Lys	Pro 5	Gly	Met	Asp	Gly	Pro 10	Lys	Val	Lys	Gln	Trp 15	Pro
Leu	Thr	Glu	Glu 20	Lys	Ile	Lys	Ala	Leu 25	Val	Glu	Ile	Cys	Thr 30	Glu	Met
Glu	Lys	Glu 35	Gly	Lys	Ile	Ser	Lys 40	Ile	Gly	Pro	Glu	Asn 45	Pro	Tyr	Asn
Thr	Pro 50	Val	Phe	Ala	Ile	Lys 55	Lys	Lys	Asp	Ser	Thr 60	Lys	Trp	Arg	Lys
Leu 65	Val	Asp	Phe	Arg	Glu 70	Leu	Asn	Lys	Arg	Thr 75	Gln	Asp	Phe	Trp	Glu 80
Val	Gln	Leu	Gly	Ile 85	Pro	His	Pro	Ala	Gly 90	Leu	Lys	Lys	Lys	Lys 95	Ser
Val	Thr	Val	Leu 100	Asp	Val	Gly	Asp	Ala 105	Tyr	Phe	Ser	Val	Pro 110	Leu	Asp
Glu	Asp	Phe 115	Arg	Lys	Tyr	Thr	Ala 120	Phe	Thr	Ile	Pro	Ser 125	Ile	Asn	Asn
Glu	Thr 130	Pro	Gly	Ile	Arg	Tyr 135	Gln	Tyr	Asn	Val	Leu 140	Pro	Gln	Gly	Trp
Lys 145	Gly	Ser	Pro	Ala	Ile 150	Phe	Gln	Ser	Ser	Met 155	Thr	Lys	Ile	Leu	Glu 160

**SECRET**







Gln	Phe	Gln	Pro	Tyr	Phe	Ala	Phe	Thr	Val	Pro	Gln	Gln	Cys	Asn	Tyr
		115					120					125			
Gly	Pro	Gly	Thr	Arg	Tyr	Ala	Trp	Lys	Val	Leu	Pro	Gln	Gly	Phe	Lys
	130					135					140				
Asn	Ser	Pro	Thr	Leu	Phe	Glu	Met	Gln	Leu	Ala	His	Ile	Leu	Gln	Pro
145					150					155					160
Ile	Arg	Gln	Ala	Phe	Pro	Gln	Cys	Thr	Ile	Leu	Gln	Tyr	Met	Asp	Asp
				165					170					175	
Ile	Leu	Leu	Ala	Ser	Pro	Ser	His	Glu	Asp	Leu	Leu	Leu	Leu	Ser	Glu
			180					185						190	
Ala	Thr	Met	Ala	Ser	Leu	Ile	Ser	His	Gly	Leu	Pro	Val	Ser	Glu	Asn
		195					200					205			
Lys	Thr	Gln	Gln	Thr	Pro	Gly	Thr	Ile	Lys	Phe	Leu	Gly	Gln	Ile	Ile
	210					215					220				
Ser	Pro	Asn	His	Leu	Thr	Tyr	Asp	Ala	Val	Pro	Thr	Val	Pro	Ile	Arg
225					230					235					240
Ser	Arg	Trp	Ala	Leu	Pro	Glu	Leu	Gln	Ala	Leu	Leu	Gly	Glu	Ile	Gln
				245					250					255	
Trp	Val	Ser	Lys	Gly	Thr	Pro	Thr	Leu	Arg	Gln	Pro	Leu	His	Ser	Leu
			260					265					270		
Tyr	Cys	Ala	Leu	Gln	Arg	His	Thr	Asp	Pro	Arg	Asp	Gln	Ile	Tyr	Leu
		275					280					285			
Asn	Pro	Ser	Gln	Val	Gln	Ser	Leu	Val	Gln	Leu	Arg	Gln	Ala	Leu	Ser
	290					295					300				
Gln	Asn	Cys	Arg	Ser	Arg	Leu	Val	Gln	Thr	Leu	Pro	Leu	Leu	Gly	Ala
305					310					315					320
Ile	Met	Leu	Thr	Leu	Thr	Gly	Thr	Thr	Thr	Val	Val	Phe	Gln	Ser	Lys
				325				330						335	
Glu	Gln	Trp	Pro	Leu	Val	Trp	Leu	His	Ala	Pro	Leu	Pro	His	Thr	Ser
			340					345					350		
Gln	Cys	Pro	Trp	Gly	Gln	Leu	Leu	Ala	Ser	Ala	Val	Leu	Leu	Leu	Asp
		355					360					365			
Lys	Tyr	Thr	Leu	Gln	Ser	Tyr	Gly	Leu	Leu	Cys	Gln	Thr	Ile	His	His
	370					375					380				
Asn	Ile	Ser	Thr	Gln	Thr	Phe	Asn	Gln	Phe	Ile	Gln	Thr	Ser	Asp	His
385					390					395					400

RECEIVED - DECEMBER 1967

Pro Ser Val Pro Ile Leu Leu His His Ser His Arg Phe Lys Asn Leu  
405 410 415

Gly Ala Gln Thr Gly Glu Leu Trp Asn Thr Phe Leu Lys Thr Ala Ala  
420 425 430

Pro Leu Ala Pro Val Lys Ala Leu Met Pro Val Phe Thr Leu Ser Pro  
435 440 445

Val Ile Ile Asn Thr Ala Pro Cys Leu Phe Ser Asp Gly Ser Thr Ser  
450 455 460

Arg Ala Ala Tyr Ile Leu Trp Asp Lys Gln Ile Leu Ser Gln Arg Ser  
465 470 475 480

Phe Pro Leu Pro Pro Pro His Lys Ser Ala Gln Arg Ala Glu Leu Leu  
485 490 495

Gly Leu Leu His Gly Leu Ser Ser Ala Arg Ser Trp Arg Cys Leu Asn  
500 505 510

Ile Phe Leu Asp Ser Lys Tyr Leu Tyr His Tyr Leu Arg Thr Leu Ala  
515 520 525

Leu Gly Thr Phe Gln Gly Arg Ser Ser Gln Ala Pro Phe Gln Ala Leu  
530 535 540

Leu Pro Arg Leu Leu Ser Arg Lys Val Val Tyr Leu His His Val Arg  
545 550 555 560

Ser His Thr Asn Leu Pro Asp Pro Ile Ser Arg Leu Asn Ala Leu Thr  
565 570 575

Asp Ala

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 555 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Asn Val Leu Tyr Arg Ile Gly Ser Asp Asn Gln Tyr Thr Gln Phe Thr  
1 5 10 15

Ile Pro Lys Lys Gly Lys Gly Val Arg Thr Ile Ser Ala Pro Thr Asp  
20 25 30

26E06031.030397

Arg	Leu	Lys	Asp	Ile	Gln	Arg	Arg	Ile	Cys	Asp	Leu	Leu	Ser	Asp	Cys
		35					40					45			
Arg	Asp	Glu	Ile	Phe	Ala	Ile	Arg	Lys	Ile	Ser	Asn	Asn	Tyr	Ser	Phe
	50					55					60				
Gly	Phe	Glu	Arg	Gly	Lys	Ser	Ile	Ile	Leu	Asn	Ala	Tyr	Lys	His	Arg
65					70					75					80
Gly	Lys	Gln	Ile	Ile	Leu	Asn	Ile	Asp	Leu	Lys	Asp	Phe	Phe	Glu	Ser
				85					90					95	
Phe	Asn	Phe	Gly	Arg	Val	Arg	Gly	Tyr	Phe	Leu	Ser	Asn	Gln	Asp	Phe
			100					105					110		
Leu	Leu	Asn	Pro	Val	Val	Ala	Thr	Thr	Leu	Ala	Lys	Ala	Ala	Cys	Tyr
		115					120					125			
Asn	Gly	Thr	Leu	Pro	Gln	Gly	Ser	Pro	Cys	Ser	Pro	Ile	Ile	Ser	Asn
	130					135					140				
Leu	Ile	Cys	Asn	Ile	Met	Asp	Met	Arg	Leu	Ala	Lys	Leu	Ala	Lys	Lys
145					150					155					160
Tyr	Gly	Cys	Thr	Tyr	Ser	Arg	Tyr	Ala	Asp	Asp	Ile	Thr	Ile	Ser	Thr
				165					170					175	
Asn	Lys	Asn	Thr	Phe	Pro	Leu	Glu	Met	Ala	Thr	Val	Gln	Pro	Glu	Gly
			180					185					190		
Val	Val	Leu	Gly	Lys	Val	Leu	Val	Lys	Glu	Ile	Glu	Asn	Ser	Gly	Phe
		195					200					205			
Glu	Ile	Asn	Asp	Ser	Lys	Thr	Arg	Leu	Thr	Tyr	Lys	Thr	Ser	Arg	Gln
	210					215					220				
Glu	Val	Thr	Gly	Leu	Thr	Val	Asn	Arg	Ile	Val	Asn	Ile	Asp	Arg	Cys
225					230					235					240
Tyr	Tyr	Lys	Lys	Thr	Arg	Ala	Leu	Ala	His	Ala	Leu	Tyr	Arg	Thr	Gly
				245					250					255	
Glu	Tyr	Lys	Val	Pro	Asp	Glu	Asn	Gly	Val	Leu	Val	Ser	Gly	Gly	Leu
			260					265					270		
Asp	Lys	Leu	Glu	Gly	Met	Phe	Gly	Phe	Ile	Asp	Gln	Val	Asp	Lys	Phe
		275					280					285			
Asn	Asn	Ile	Lys	Lys	Lys	Leu	Asn	Lys	Gln	Pro	Asp	Arg	Tyr	Val	Leu
	290					295					300				
Thr	Asn	Ala	Thr	Leu	His	Gly	Phe	Lys	Leu	Lys	Leu	Asn	Ala	Arg	Glu
305					310					315					320

**SECRET**

030603031 030397

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Arg	Trp	Phe	Ser	Phe	His	Arg	Glu	Val	Asp	Thr	Gly	Thr	His	Tyr	Gln
1				5					10					15	
Thr	Trp	Glu	Ile	Pro	Lys	Arg	Asp	Gly	Gly	Lys	Arg	Thr	Leu	Thr	Ala
			20					25					30		
Pro	Lys	Arg	Glu	Leu	Lys	Ala	Val	Gln	Arg	Trp	Val	Leu	Ala	Asn	Val
		35					40					45			
Val	Glu	Arg	Leu	Pro	Val	His	Gly	Ala	Ala	His	Gly	Phe	Val	Ala	Gly
	50					55					60				
Arg	Ser	Ile	Leu	Thr	Asn	Ala	Leu	Ala	His	Gln	Gly	Ala	Asp	Val	Val
65					70					75					80
Val	Lys	Val	Asp	Met	Lys	Asp	Phe	Phe	Pro	Ser	Val	Thr	Trp	Pro	Arg
				85					90					95	
Val	Lys	Gly	Leu	Leu	Arg	Lys	Gly	Gly	Leu	Pro	Glu	Asn	Leu	Ala	Thr
			100					105					110		
Leu	Leu	Ala	Leu	Leu	Ser	Thr	Glu	Ala	Pro	Arg	Glu	Val	Val	Arg	Phe
		115					120					125			
Arg	Gly	Glu	Thr	Leu	Tyr	Val	Ala	Lys	Gly	Pro	Arg	Ala	Leu	Pro	Gln
	130					135						140			
Gly	Ala	Pro	Thr	Ser	Pro	Ala	Leu	Thr	Asn	Ala	Leu	Cys	Leu	Arg	Leu
145					150					155					160
Asp	Lys	Arg	Leu	Ser	Ala	Leu	Ser	Lys	Arg	Leu	Gly	Phe	Thr	Tyr	Thr
				165					170					175	
Arg	Tyr	Ala	Asp	Asp	Leu	Thr	Phe	Ser	Trp	Arg	Arg	Ala	Lys	Lys	Ser
			180					185					190		
Arg	Gln	Lys	Glu	Leu	Pro	Leu	Ala	Asp	Ala	Pro	Val	Ala	Leu	Leu	Leu
		195					200					205			
Ala	Arg	Val	Lys	Gly	Val	Leu	Glu	Ala	Glu	Gly	Phe	Thr	Leu	His	Pro
	210					215					220				
Asp	Lys	Thr	Arg	Val	Gln	Arg	Lys	Gly	Ser	Arg	Gln	Arg	Val	Thr	Gly
225					230					235					240
Leu	Val	Val													

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

ORIGINAL-030397

00000001 = 030397

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Arg 1	Trp	Phe	Ala	Phe 5	His	Arg	Glu	Val	Asp 10	Thr	Ala	Thr	His	Tyr 15	Val
Ser	Trp	Thr	Ile 20	Pro	Lys	Arg	Asp	Gly 25	Ser	Lys	Arg	Thr	Ile 30	Thr	Ser
Pro	Lys	Pro 35	Glu	Leu	Lys	Ala	Ala 40	Gln	Arg	Trp	Val	Leu 45	Ser	Asn	Val
Val	Glu 50	Arg	Leu	Pro	Val	His 55	Gly	Ala	Ala	His	Gly 60	Phe	Val	Ala	Gly
Arg 65	Ser	Ile	Leu	Thr	Asn 70	Ala	Leu	Ala	His	Gln 75	Gly	Ala	Asp	Val	Val 80
Val	Lys	Val	Asp	Leu 85	Lys	Asp	Phe	Phe	Pro 90	Ser	Val	Thr	Trp	Arg 95	Arg
Val	Lys	Gly	Leu 100	Leu	Arg	Lys	Gly	Gly 105	Leu	Arg	Glu	Gly	Thr 110	Ser	Thr
Leu	Leu	Ser 115	Leu	Leu	Ser	Thr	Glu 120	Ala	Pro	Arg	Glu	Ala 125	Val	Gln	Phe
Pro	Arg 130	Glu	Leu	Leu	His	Val 135	Ala	Lys	Gly	Pro	Arg 140	Ala	Leu	Pro	Gln
Gly 145	Ala	Pro	Thr	Ser	Pro 150	Gly	Ile	Thr	Asn	Ala 155	Leu	Cys	Leu	Lys	Leu 160
Asp	Lys	Arg	Leu	Ser 165	Ala	Leu	Ala	Lys	Arg 170	Leu	Gly	Phe	Thr	Tyr 175	Thr
Arg	Tyr	Ala	Asp 180	Asp	Leu	Thr	Phe	Ser 185	Trp	Thr	Lys	Ala	Lys 190	Gln	Pro
Lys	Pro	Arg 195	Arg	Thr	Gln	Arg	Pro 200	Pro	Val	Ala	Val	Leu 205	Leu	Ser	Arg
Val	Gln 210	Glu	Val	Val	Glu	Ala 215	Glu	Gly	Phe	Arg	Val 220	His	Pro	Asp	Lys
Thr 225	Arg	Val	Ala	Arg	Lys 230	Gly	Thr	Arg	Gln	Arg 235	Val	Thr	Gly	Leu	Val 240

Val

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 231 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Arg His Tyr Ser Ile His Arg Pro Arg Glu Arg Val Arg His Tyr Val  
1 5 10 15  
Thr Phe Ala Val Pro Lys Arg Ser Gly Gly Val Arg Leu Leu His Ala  
20 25 30  
Pro Lys Arg Arg Leu Lys Ala Leu Gln Arg Arg Met Leu Ala Leu Leu  
35 40 45  
Val Ser Lys Leu Pro Val Ser Pro Gln Ala His Gly Phe Val Pro Gly  
50 55 60  
Arg Ser Ile Lys Thr Gly Ala Ala Pro His Val Gly Arg Arg Val Val  
65 70 75 80  
Leu Lys Leu Asp Leu Lys Asp Phe Phe Pro Ser Val Thr Phe Ala Arg  
85 90 95  
Val Arg Gly Leu Leu Lys Ala Leu Gly Tyr Gly Tyr Pro Val Ala Ala  
100 105 110  
Thr Leu Ala Val Leu Met Thr Glu Ser Glu Arg Gln Pro Val Glu Leu  
115 120 125  
Glu Gly Ile Leu Phe His Val Pro Val Gly Pro Arg Val Cys Val Gln  
130 135 140  
Gly Ala Pro Thr Ser Pro Ala Leu Cys Asn Ala Val Leu Leu Arg Leu  
145 150 155 160  
Asp Arg Arg Leu Ala Gly Leu Ala Arg Arg Tyr Gly Tyr Thr Tyr Thr  
165 170 175  
Arg Tyr Ala Asp Asp Leu Thr Phe Ser Gly Asp Asp Val Thr Ala Leu  
180 185 190  
Glu Arg Val Arg Ala Leu Ala Ala Arg Tyr Val Gln Glu Glu Gly Phe  
195 200 205

0808031.030397



Glu Val Asn Arg Glu Lys Thr Arg Val Gln Arg Arg Gly Gly Ala Gln  
 210 215 220

Arg Val Thr Gly Val Thr Val  
 225 230

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 234 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Phe Leu Thr Asn Val Leu Tyr Arg Ile Gly Ser Asp Asn Gln Tyr Thr  
 1 5 10 15  
 Gln Phe Thr Ile Pro Lys Lys Gly Lys Gly Val Arg Thr Ile Ser Ala  
 20 25 30  
 Pro Thr Asp Arg Leu Lys Asp Ile Gln Arg Arg Ile Cys Asp Leu Leu  
 35 40 45  
 Ser Asp Cys Arg Asp Glu Ile Phe Ala Ile Arg Lys Ile Ser Asn Asn  
 50 55 60  
 Tyr Ser Phe Gly Phe Glu Arg Gly Lys Ser Ile Ile Leu Asn Ala Tyr  
 65 70 75 80  
 Lys His Arg Gly Lys Gln Ile Ile Leu Asn Ile Asp Leu Lys Asp Phe  
 85 90 95  
 Phe Glu Ser Phe Asn Phe Gly Arg Val Arg Gly Tyr Phe Leu Ser Asn  
 100 105 110  
 Gln Asp Phe Leu Leu Asn Pro Val Val Ala Thr Thr Leu Ala Lys Ala  
 115 120 125  
 Ala Cys Tyr Asn Gly Thr Leu Pro Gln Gly Ser Pro Cys Ser Pro Ile  
 130 135 140  
 Ile Ser Asn Leu Ile Cys Asn Ile Met Asp Met Arg Leu Ala Lys Leu  
 145 150 155 160  
 Ala Lys Lys Tyr Gly Cys Thr Tyr Ser Arg Tyr Ala Asp Asp Ile Thr  
 165 170 175  
 Ile Ser Thr Asn Lys Asn Thr Phe Pro Leu Glu Met Ala Thr Val Gln  
 180 185 190

26E0E0-TE080840

Pro Glu Gly Val Val Leu Gly Lys Val Leu Val Lys Glu Ile Glu Asn  
 195 200 205

Ser Gly Phe Glu Ile Asn Asp Ser Lys Thr Arg Leu Thr Tyr Lys Thr  
 210 215 220

Ser Arg Gln Glu Val Thr Gly Leu Thr Val  
 225 230

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 215 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Val Glu Thr Leu Arg Leu Leu Ile Tyr Thr Ala Asp Phe Arg Tyr Arg  
 1 5 10 15

Ile Tyr Thr Val Glu Lys Lys Gly Pro Glu Lys Arg Met Arg Thr Ile  
 20 25 30

Tyr Gln Pro Ser Arg Glu Leu Lys Ala Leu Gln Gly Trp Val Leu Arg  
 35 40 45

Asn Ile Leu Asp Lys Leu Ser Ser Ser Pro Phe Ser Ile Gly Phe Glu  
 50 55 60

Lys His Gln Ser Ile Leu Asn Asn Ala Thr Pro His Ile Gly Ala Asn  
 65 70 75 80

Phe Ile Leu Asn Ile Asp Leu Glu Asp Phe Phe Pro Ser Leu Thr Ala  
 85 90 95

Asn Lys Val Phe Gly Val Phe His Ser Leu Gly Tyr Asn Arg Leu Ile  
 100 105 110

Ser Ser Val Leu Thr Lys Ile Cys Cys Tyr Lys Asn Leu Leu Pro Gln  
 115 120 125

Gly Ala Pro Ser Ser Pro Lys Leu Ala Asn Leu Ile Cys Ser Lys Leu  
 130 135 140

Asp Tyr Arg Ile Gln Gly Tyr Ala Gly Ser Arg Gly Leu Ile Tyr Thr  
 145 150 155 160

Arg Tyr Ala Asp Asp Leu Thr Leu Ser Ala Gln Ser Met Lys Lys Val  
 165 170 175

46E0E0-TE080880



Asp Ile Ile Val Ser Thr Asn Met Lys Gly Ala Ser Lys Leu Ile Leu  
180 185 190

Asp Cys Phe Lys Arg Thr Met Lys Glu Ile Gly Pro Asp Phe Lys Ile  
195 200 205

Asn Ile Lys Lys Phe Lys Ile Cys Ser Ala Ser Gly Gly Ser Ile Val  
210 215 220

Val Thr Gly Leu Lys Val  
225 230

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Ile Gln Arg Leu His Ala Leu Ser Asn His Ala Gly Arg His Tyr Arg  
1 5 10 15

Arg Ile Ile Leu Ser Lys Arg His Gly Gly Gln Arg Leu Val Leu Ala  
20 25 30

Pro Asp Tyr Leu Leu Lys Thr Val Gln Arg Asn Ile Leu Lys Asn Val  
35 40 45

Leu Ser Gln Phe Pro Leu Ser Pro Phe Ala Thr Ala Tyr Arg Pro Gly  
50 55 60

Cys Pro Ile Val Ser Asn Ala Gln Pro His Cys Gln Gln Pro Gln Ile  
65 70 75 80

Leu Lys Leu Asp Ile Glu Asn Phe Phe Asp Ser Ile Ser Trp Leu Gln  
85 90 95

Val Trp Arg Val Phe Arg Gln Ala Gln Leu Pro Arg Asn Val Val Thr  
100 105 110

Met Leu Thr Trp Ile Cys Cys Tyr Asn Asp Ala Leu Pro Gln Gly Ala  
115 120 125

Pro Thr Ser Pro Ala Ile Ser Asn Leu Val Met Arg Arg Phe Asp Glu  
130 135 140

Arg Ile Gly Glu Trp Cys Gln Ala Arg Gly Ile Thr Tyr Thr Arg Tyr  
145 150 155 160

J6558031.030397

55

60

65

CGC Arg 70	CGC Arg	TAC Tyr	ACC Thr	CCG Pro	GGC Gly 75	CGG Arg	AAG Lys	AAG Lys	TGG Trp	ATG Met 80	GAG Glu	GCC Ala	GCC Ala	GAG Glu	GCC Ala 85	533
CGG Arg	CGG Arg	CTG Leu	TTC Phe	TCC Ser 90	GCC Ala	ACG Thr	CTG Leu	CGC Arg	ACG Thr 95	CGG Arg	AAC Asn	CGG Arg	AAC Asn	CTG Leu 100	AGG Arg	581
GAC Asp	TTG Leu	CTG Leu	CCC Pro 105	GAC Asp	GAG Glu	GCA Ala	CAG Gln	CTG Leu 110	GCG Ala	CGC Arg	TAC Tyr	GGC Gly	CTG Leu 115	CCG Pro	GTC Val	629
TGG Trp	CGC Arg	ACG Thr 120	GAA Glu	GAG Glu	GAC Asp	GTG Val	GCA Ala	GCG Ala	GCC Ala	CTG Leu	GGC Gly	GTC Val 130	TCG Ser	GTG Val	GGC Gly	677
GTG Val 135	CTC Leu	CGC Arg	CAC His	TAC Tyr	AGC Ser	ATC Ile 140	CAC His	CGC Arg	CCG Pro	CGC Arg	GAG Glu 145	CGG Arg	GTG Val	CGG Arg	CAC His	725
TAC Tyr 150	GTG Val	ACC Thr	TTC Phe	GCC Ala	GTG Val 155	CCC Pro	AAG Lys	CGC Arg	TCC Ser	GGA Gly 160	GGC Gly	GTC Val	CGG Arg	CTG Leu	CTG Leu 165	773
CAT His	GCG Ala	CCC Pro	AAG Lys	CGG Arg 170	CGC Arg	CTG Leu	AAG Lys	GCC Ala	CTG Leu 175	CAA Gln	CGC Arg	CGG Arg	ATG Met	CTG Leu 180	GCG Ala	821
CTC Leu	CTG Leu	GTG Val	TCG Ser 185	AAG Lys	CTC Leu	CCC Pro	GTG Val	AGT Ser 190	CCA Pro	CAG Gln	GCC Ala	CAT His	GGC Gly 195	TTC Phe	GTG Val	869
CCC Pro 200	GGC Gly	CGC Arg	TCC Ser	ATC Ile	AAG Lys	ACG Thr	GGC Gly 205	GCC Ala	GCG Ala	CCG Pro	CAC His	GTG Val 210	GGC Gly	CGG Arg	CGG Arg	917
GTG Val 215	GTC Val	CTG Leu	AAG Lys	CTG Leu	GAC Asp	CTG Leu 220	AAG Lys	GAC Asp	TTC Phe	TTC Phe	CCC Pro 225	TCC Ser	GTC Val	ACC Thr	TTC Phe	965
GCG Ala 230	CGG Arg	GTG Val	CGA Arg	GGG Gly	CTG Leu 235	CTC Leu	ATC Ile	GCC Ala	CTG Leu	GGC Gly 240	TAC Tyr	GGC Gly	TAT Tyr	CCC Pro	GTG Val 245	1013
GCG Ala	GCC Ala	ACG Thr	CTC Leu	GCG Ala 250	GTG Val	CTG Leu	ATG Met	ACG Thr	GAG Glu 255	TCC Ser	GAG Glu	CGC Arg	CAG Gln	CCC Pro 260	GTG Val	1061
GAG Glu	CTG Leu	GAG Glu	GGC Gly 265	ATC Ile	CTC Leu	TTC Phe	CAC His	GTT Val 270	CCC Pro	GTG Val	GGC Gly	CCA Pro	CGC Arg 275	GTC Val	TGC Cys	1109

cy  
03308031-03009  
460000-TE008000

CTC  
GCC  
GGC  
TCG

(i) SEQUENCE CHARACTERISTICS:

- (ix) FEATURE:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CTCCGAGCCC	GCCTCCGAGG	ACGCGCTCGC	GGCCCCGGGCG	GCGGGGGCGG	ACGCGCGGCCG		60
GCGGCCCCACG	GAGACGCTTG	ACCCGGGAGA	CGACGAATGA	CGATAACGGC	AGGTGCTCTC		120
GGGAGAGGCC	AGGGCTCGCA	GATGAGCCAT	GAGTACCGCG	GTGTTTCGCC	GCGGGGGTGT		180
TCTGTCCCCA	TCTCTTCGCC	AGGGTCCCAG	CGTACGCAAC	GCAGGGAGCC	CCGGGTCCAA		240
CGCCTCGCAG	GTCGTCCCCT	GGCCTCTTCC	GGAGCACC	ATG AGC TGG TTC GAC			293
				Met Ser Trp Phe Asp			
				1	5		
ACC ACC CTC TCC CGG CTC AAG GGG TTG TTC AGC CGT CCC GTG ACA CGA							341
Thr Thr Leu Ser Arg Leu Lys Gly Leu Phe Ser Arg Pro Val Thr Arg							
			10			15	
AGC ACC ACC GGG CTG GAC GTG CCG CTG GAT GCC CAC GGA CGT CCC CAG							389
Ser Thr Thr Gly Leu Asp Val Pro Leu Asp Ala His Gly Arg Pro Gln							
			25			30	
GAC GTC GTG ACG GAG ACG GTC TCC ACG TCG GGC CCC CTG AAG CCA GGG							437
Asp Val Val Thr Glu Thr Val Ser Thr Ser Gly Pro Leu Lys Pro Gly							
			40			45	
CAC CTG CGA CAG GTC CGC CGG GAT GCG CGG CTG CTC CCC AAG GGC GTC							485
His Leu Arg Gln Val Arg Arg Asp Ala Arg Leu Leu Pro Lys Gly Val							

GTG CAG GGC GCC CCC ACG AGC CCC GCC CTG TGC AAC GCG GTG CTG CTG 1157  
Val Gln Gly Ala Pro Thr Ser Pro Ala Leu Cys Asn Ala Val Leu Leu  
280 285 290

CGA CTG GAC CGG CGG CTG GCG GGA CTG GCG CGT CGG TAC GGC TAC ACG 1205  
Arg Leu Asp Arg Arg Leu Ala Gly Leu Ala Arg Arg Tyr Gly Tyr Thr  
295 300 305

TAC ACG CGC TAC GCG GAT GAC CTC ACC TTC TCC GGC GAC GAC GTC ACG 1253  
Tyr Thr Arg Tyr Ala Asp Asp Leu Thr Phe Ser Gly Asp Asp Val Thr  
310 315 320 325

GCG CTG GAG CGA GTC CGC GCG CTG GCC GCG CGG TAC GTG CAG GAG GAA 1301  
Ala Leu Glu Arg Val Arg Ala Leu Ala Ala Arg Tyr Val Gln Glu Glu  
330 335 340

GGC TTC GAG GTC AAC CGC GAG AAG ACC CGC GTG CAG CGC CGG GGC GGT 1349  
Gly Phe Glu Val Asn Arg Glu Lys Thr Arg Val Gln Arg Arg Gly Gly  
345 350 355

GCC CAG CGC GTC ACT GGC GTC ACC GTG AAT ACG ACG CTG GGC TTG TCA 1397  
Ala Gln Arg Val Thr Gly Val Thr Val Asn Thr Thr Leu Gly Leu Ser  
360 365 370

CGC GAG GAG CGG CCG CGG CTC CGG GCG ATG CTG CAC CAG GAG GCG CGG 1445  
Arg Glu Glu Arg Pro Arg Leu Arg Ala Met Leu His Gln Glu Ala Arg  
375 380 385

TCG GAG GAC GTC GAG GCA CAC CGC GCG CAC CTC GAC GGC CTC CTG GCC 1493  
Ser Glu Asp Val Glu Ala His Arg Ala His Leu Asp Gly Leu Leu Ala  
390 395 400 405

TAC GTG AAG ATG CTC AAC CCG GAG CAG GCG GAG CGG CTC GCT CGC CGG 1541  
Tyr Val Lys Met Leu Asn Pro Glu Gln Ala Glu Arg Leu Ala Arg Arg  
410 415 420

CGC AAG CCG CGC GGG ACG TGAGCGAGGG CTCAGCTCCG GATGGGCCAG 1589  
Arg Lys Pro Arg Gly Thr  
425

GGCCTGTCAC GCGTCCCGGC CTCCCAGTTG TCATGGCGGC CGTCCCAGTA C 1640

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3060 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS

(B) LOCATION: 763..2202

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CCCACTTCCG GCGCTCGGGC TGC GCGAGGG CCCGTGCGAG CACATGATGG CGCTGCGGCT 60  
CGTCCAGGTC CGGCACCGCG CCGAGCAGGA AGCACTGCGT CAGACCCCCG CGGGCCGCCA 120  
GCTCATCCGC GCGGAGACGC GCTCCTACGT GCGGCGCGAG CCCTCCGGCC AGGAGCAGGT 180  
GTACCGCGTC TCATTGGATG GGAAAGTGGT GGCGGTGGAG TGGGGCCCCC GCCAGGGGGA 240  
GTCCCGCCGG CAGAAGCTCT GGTTCGACAC GGACGCCGAG GCGCGCACCG CCTACTTCAC 300  
GCGCCTGGAG TCCTTGGCCG CGGAGGGATA TATCGATGCG GCTGCTTCAA TGATGTAGAA 360  
CACGCAAGCC ACGGGGCCGC GGGCGCGCGG CGGAAAGGCA GGTGCGACGG AACGACAGAC 420  
ACTCGTGCGA GCGACCGAGA GAGGTCCCAA GCCATCAGCC TCAGCGCCTC GAGCGCGAGA 480  
GCGGCGTTGC GCCGCTCTGG TTGAATTGCA GGACACTCTC CGCAAGGTAG CCTGTTCTTG 540  
GCTCTCTTCC CTCCGGTGAG TACCTCTCCG GCCGGGGAGC TGAACCAACG ACGCAACCGC 600  
CGTTTCCCCG GCCGGAGAGG TACTCACCGG AGGGGAGAGC CGGTGAGGCT ACCGTGCCCC 660  
AGGTGAGAAG GTGGTGCCTT CGGGCCTCCC TCGACCGCTC GCGCTCCGTC GCCCTGCCCT 720  
GCCTCGCCCC CCCACCTTG CTCACCGGCG CCAGGAGCCG TC ATG ACC GCC AAG 774  
Met Thr Ala Lys  
1  
CTG GAG TCA CAC GTC CCC GCC GCG CCC CCC GTC TCC GCC GAG GCG CCC 822  
Leu Glu Ser His Val Pro Ala Ala Pro Pro Val Ser Ala Glu Ala Pro  
5 10 15 20  
GCC CCC ACC CGT CCC GAT GCC GCG AAG CAG GAG GCC CGC CGC GCC CAC 870  
Ala Pro Thr Arg Pro Asp Ala Ala Lys Gln Glu Ala Arg Arg Ala His  
25 30 35  
CAC GAG GCG CTG CGC CTG CGG TGG AAG GCC ATC GAA GAG GCG GGC GGC 918  
His Glu Ala Leu Arg Leu Arg Trp Lys Ala Ile Glu Glu Ala Gly Gly  
40 45 50  
ACG GAC GCC TGG GTG CGG CAG CAG CTG GTG GCC AAG GGC GTC GCG GCG 966  
Thr Asp Ala Trp Val Arg Gln Gln Leu Val Ala Lys Gly Val Ala Ala  
55 60 65  
GAA GAG GTG GAC TTC GAG TCG CTC AGC GAC AAG CAG AAG GCG GCC TGG 1014  
Glu Glu Val Asp Phe Glu Ser Leu Ser Asp Lys Gln Lys Ala Ala Trp  
70 75 80  
AAG GAG AAG AAG AAG GCC GAG GCC ACC GAG CGG CGC GCG CAG AAG CGC 1062



Lys 85	Glu	Lys	Lys	Lys	Ala 90	Glu	Ala	Thr	Glu	Arg 95	Arg	Ala	Gln	Lys	Arg 100	
CTG Leu	GCG Ala	TGG Trp	GAG Glu	GCC Ala 105	TGG Trp	AAG Lys	GCC Ala	ACG Thr	CAC His 110	ATC Ile	CAC His	CAC His	CTG Leu	GGC Gly 115	GTG Val	1110
GGG Gly	GTG Val	CAC His	TGG Trp 120	GAC Asp	GAG Glu	GCC Ala	GGA Gly	GGG Gly 125	CCG Pro	GAC Asp	AAG Lys	TTC Phe	GAC Asp 130	GTG Val	GCC Ala	1158
GGG Gly	CGC Arg	GAG Glu 135	GAG Glu	CGG Arg	GCC Ala	AAG Lys	GCC Ala 140	AAC Asn	GGC Gly	TTG Leu	CCG Pro	GAG Glu 145	GGG Gly	TTG Leu	GAC Asp	1206
TCG Ser	GTC Val 150	GAG Glu	GCG Ala	CTG Leu	GCC Ala	AAA Lys 155	GCG Ala	CTG Leu	GGC Gly	ATC Ile	TCC Ser 160	GTG Val	TCG Ser	CGC Arg	CTG Leu	1254
CGC Arg 165	TGG Trp	TTC Phe	TCC Ser	TTC Phe	CAC His 170	CGC Arg	GAG Glu	GTG Val	GAC Asp	ACG Thr 175	GGC Gly	ACG Thr	CAC His	TAC Tyr	CAG Gln 180	1302
ACG Thr	TGG Trp	GAG Glu	ATT Ile	CCG Pro 185	AAG Lys	CGG Arg	GAC Asp	GGC Gly 190	GGC Gly	AAG Lys	CGG Arg	ACG Thr	CTC Leu	ACC Thr 195	GCG Ala	1350
CCG Pro	AAG Lys	CGG Arg	GAG Glu 200	CTC Leu	AAG Lys	GCC Ala	GTG Val	CAG Gln 205	CGC Arg	TGG Trp	GTG Val	CTC Leu	GCG Ala 210	AAC Asn	GTG Val	1398
GTG Val	GAG Glu	CGG Arg 215	CTG Leu	CCG Pro	GTG Val	CAC His	GGG Gly 220	GCC Ala	GCG Ala	CAC His	GGC Gly	TTC Phe 225	GTG Val	GCG Ala	GGG Gly	1446
CGC Arg	TCC Ser 230	ATC Ile	CTC Leu	ACC Thr	AAC Asn	GCG Ala 235	CTG Leu	GCC Ala	CAC His	CAG Gln	GGC Gly 240	GCG Ala	GAC Asp	GTG Val	GTG Val	1494
GTG Val 245	AAG Lys	GTG Val	GAC Asp	ATG Met	AAG Lys 250	GAC Asp	TTC Phe	TTC Phe	CCT Pro	TCC Ser 255	GTG Val	ACG Thr	TGG Trp	CCC Pro	CGG Arg 260	1542
GTC Val	AAG Lys	GGA Gly	CTG Leu	CTG Leu 265	CGC Arg	AAG Lys	GGA Gly	GGA Gly	CTC Leu 270	CCG Pro	GAG Glu	AAC Asn	CTG Leu	GCG Ala 275	ACG Thr	1590
CTC Leu	CTG Leu	GCG Ala	CTG Leu 280	CTC Leu	TCC Ser	ACC Thr	GAG Glu	GCC Ala 285	CCG Pro	CGC Arg	GAG Glu	GTG Val	GTG Val 290	CGG Arg	TTC Phe	1638
CGG Arg	GGA Gly	GAG Glu 295	ACG Thr	CTG Leu	TAC Tyr	GTG Val	GCC Ala 300	AAG Lys	GGC Gly	CCT Pro	CGC Arg	GCG Ala 305	CTG Leu	CCC Pro	CAG Gln	1686

GGG Gly	GCC Ala	CCC Pro	ACC Thr	TCT Ser	CCG Pro	GCG Ala	CTG Leu	ACG Thr	AAC Asn	GCG Ala	CTG Leu	TGC Cys	CTG Leu	CGG Arg	CTG Leu	1734
310 315 320																
GAC Asp	AAG Lys	CGG Arg	CTC Leu	TCG Ser	GCG Ala	CTG Leu	TCG Ser	AAG Lys	CGG Arg	CTG Leu	GGC Gly	TTC Phe	ACG Thr	TAC Tyr	ACG Thr	1782
325 330 335 340																
CGC Arg	TAT Tyr	GCG Ala	GAT Asp	GAC Asp	CTG Leu	ACG Thr	TTC Phe	TCC Ser	TGG Trp	CGG Arg	CGG Arg	GCG Ala	AAG Lys	AAG Lys	TCC Ser	1830
345 350 355																
CGG Arg	CAG Gln	AAG Lys	GAA Glu	CTC Leu	CCC Pro	CTG Leu	GCG Ala	GAT Asp	GCG Ala	CCG Pro	GTG Val	GCG Ala	CTG Leu	CTC Leu	CTG Leu	1878
360 365 370																
GCG Ala	CGG Arg	GTG Val	AAG Lys	GGT Gly	GTG Val	CTG Leu	GAG Glu	GCC Ala	GAG Glu	GGT Gly	TTC Phe	ACG Thr	CTG Leu	CAC His	CCG Pro	1926
375 380 385																
GAC Asp	AAG Lys	ACG Thr	CGG Arg	GTG Val	CAG Gln	CGC Arg	AAG Lys	GGC Gly	AGC Ser	CGG Arg	CAG Gln	CGG Arg	GTG Val	ACG Thr	GGG Gly	1974
390 395 400																
CTC Leu	GTG Val	GTG Val	AAC Asn	GAG Glu	GCC Ala	CCC Pro	GAG Glu	GGC Gly	GTT Val	CCG Pro	GGT Gly	GCC Ala	CGG Arg	GTG Val	CCC Pro	2022
405 410 415 420																
CGC Arg	GAT Asp	GTG Val	GTG Val	CGG Arg	CGG Arg	CTG Leu	CGC Arg	GCG Ala	GCG Ala	ATC Ile	CAC His	AAC Asn	CGG Arg	GAG Glu	CAG Gln	2070
425 430 435																
GGC Gly	AAG Lys	CCC Pro	GGC Gly	CCC Pro	ACC Thr	GGG Gly	GAG Glu	ACG Thr	CTG Leu	GAG Glu	CAG Gln	CTC Leu	AAG Lys	GGG Gly	CTC Leu	2118
440 445 450																
GCG Ala	GCC Ala	TTC Phe	CTT Leu	CAC His	ATG Met	ACG Thr	GAC Asp	GCG Ala	GAG Glu	AAG Lys	GGC Gly	CGC Arg	GCC Ala	TTC Phe	CTG Leu	2166
455 460 465																
CGA Arg	CGG Arg	CTG Leu	GAG Glu	GCC Ala	CTC Leu	GAG Glu	AAG Lys	CGC Arg	CAG Gln	ACC Thr	GCC Ala	TGACCCTCAC				2212
470 475 480																
TGGTCGTCCG GGGCATCGCA GCGGGCGCCG GGACGGACCG TCACCCCCCA GATCTCCATG 2272																
CCATGCTGGG GATTCTGGGC GGTGAAGAAG ACTTCCCAGC CGAGACGGAC GAAGCCCTGC 2332																
GGATCCGATG ACTCCTCGCC CGGGGCGATC TCCCGGAGGG GCACCGTTCC GACGTCCGTG 2392																
CCATTGCTCA CCCAGGGCTC CCGGCCCCAG CCTTGGGTGT CCGCCGAGAA GAAGAGCAGC 2452																
CCGGAGATGG CCGTCAGGTT CTCCGGCGAC GCATCCTCGG GGCCCGGCGC CAAATCCTTC 2512																

AGCAGCAGGG TGCCCTTGGC GGTGCCATCG CTGGACCACA GCTCCCGGCC GTGGAGGCTG 2572  
TCACTCGCGG CGAAGTAGAG CATCCCATTC AGCGCCTTGA TGGCGCTGGG CGCCGAGCTG 2632  
TCCGGACCCG GCCAGATGTC CTTACCCCGG ACCGTGCCAT GCGACGTGCC ATCGCTGACC 2692  
CACAGCTCCT CGCCCTCGGG CTGGCCCCAG AACTCGGGCT CGCCTCCCCC GGCGCTGAAG 2752  
AAGATCTTCC CCCCAGAGCGC CGTGAGATCA TGC GGATAGA GGCCGGGGAA GAAGCGCAGC 2812  
TGCTCGGAGA CGGTGCCTCT GGAGCACCAC AGGCTGGCCT CGCCTTCGTC ATTGTCGAGC 2872  
AGGAAGAAGA GCACCGAGTC CGCCGCGGTG AACGCGGAGA GGAAGTTGTC CTCGGGGCCC 2932  
GTGAAGACAG ACGTGGTGCT GGACAGCCCC AGGCTGCGCC AGATGAACAC CTCGTCATTG 2992  
ACGTTGGCCA CGAAGAAGAG CGCATCGCCG ACCCGGGTGA GCCGGCGCGG GCTGGAGCTG 3052  
CCGGGCAC 3060

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2788 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..103

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 707..1654

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1644..2591

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

T TTC GAG AAG CGC CAT ACC AAA CAG GGG ATA CAG ACC AAC CTG ACG 46  
Phe Glu Lys Arg His Thr Lys Gln Gly Ile Gln Thr Asn Leu Thr  
1 5 10 15  
CTG AAA GAG GAA AGC TAC GGC GAC TGG CTG CCG AAG TGC GAC GAC CCC 94  
Leu Lys Glu Glu Ser Tyr Gly Asp Trp Leu Pro Lys Cys Asp Asp Pro  
20 25 30  
GCA GCA ACA TAACCTCACT CAGACCGGCA ACAGCCGGTC TTTTCCTTTC 143

2572-3060

Ala Ala Thr

TGGCCATTGC CACAAGGTGA ACAATCCACT GTTCACCCTT CACCGTTTAT TCACCCTTTA 203  
TCACTATGAA ATTATTAATA AAAAACCAGA GGTGAACAGT GTGAACAGTA AAACCTGAAA 263  
AAACTTTTTTA TCACCCCGCG CATCGCCCGA CTGGACAGAT CCAGAACGAG CAAAAATCAC 323  
AAAGGTGACG AGTCGACTGT TCACTCTTCA CCAACTCATC ACCACCTAAC CACATGATAT 383  
AAAATGATAA ATAATCGAGG TGAACAGTTA AATGCAAAAA AACTTTTTTCT CAGCTCTTGG 443  
ATAAAAGAAA ATTAATTCAC ATCAATAGCT TTCCTCTTGA ATCCTCTTGA GGTTTATGAG 503  
AGCGTAACAG AGCCAAACCT AGCATTTTAT GGGTTAATAG CCCATCGCGC ATGAGTCATG 563  
GTTTCGCCTA GTATTTTAGC TATGCCCGTC GTTCAGTTCG CTGAGCGGCG GCTGGGGGCC 623  
ACCGATCAGC GAACTGATCG ACGTGCTCAA GTAGGTTTGG CTCTTTTAGT CCTCTACCAT 683  
CAAGGTGCAT AAGGATATTC TCG ATG CTG ACT CAG CTA AAA AAA AAT GGT 733  
Met Leu Thr Gln Leu Lys Lys Asn Gly  
1 5  
ACT GAG GTA TCT AGA GCA ACC GCG TTA TTT TCA TCA TTC GTT GAA AAG 781  
Thr Glu Val Ser Arg Ala Thr Ala Leu Phe Ser Ser Phe Val Glu Lys  
10 15 20 25  
AAC AAA GTA AAA TGT CCT GGT AAT GTA AAA AAA TTC GTC TTT CTG TGT 829  
Asn Lys Val Lys Cys Pro Gly Asn Val Lys Lys Phe Val Phe Leu Cys  
30 35 40  
GGT GCT AAC AAA AAC AAT GGA GAA CCA TCA GCA AGA CGA TTG GAA TTA 877  
Gly Ala Asn Lys Asn Asn Gly Glu Pro Ser Ala Arg Arg Leu Glu Leu  
45 50 55  
ATA AAT TTT TCT GAA AGG TAT TTG AAT AAC TGT CAC TTT TTT CTT GCT 925  
Ile Asn Phe Ser Glu Arg Tyr Leu Asn Asn Cys His Phe Phe Leu Ala  
60 65 70  
GAA CTA GTT TTC AAA GAA TTA AGC ACC GAT GAA GAA TCA TTA TCT GAT 973  
Glu Leu Val Phe Lys Glu Leu Ser Thr Asp Glu Glu Ser Leu Ser Asp  
75 80 85  
AAT TTA TTA GAT ATC GAA GCT GAC TTA TCT AAA TTA GCT GAT CAT ATT 1021  
Asn Leu Leu Asp Ile Glu Ala Asp Leu Ser Lys Leu Ala Asp His Ile  
90 95 100 105  
ATC ATT GTT TTA GAA AGT TAT TCA TCT TTC ACG GAA CTT GGT GCA TTC 1069  
Ile Ile Val Leu Glu Ser Tyr Ser Ser Phe Thr Glu Leu Gly Ala Phe  
110 115 120  
GCA TAC AGC AAG CAA TTA CGC AAG AAA TTA ATA ATA GTT AAC AAT ACA 1117

Ala	Tyr	Ser	Lys 125	Gln	Leu	Arg	Lys	Lys 130	Leu	Ile	Ile	Val	Asn 135	Asn	Thr						
AAA	TTT	ATA	AAT	GAG	AAA	TCA	TTT	ATA	AAT	ATG	GGA	CCA	ATA	AAG	GCT	1165					
Lys	Phe	Ile 140	Asn	Glu	Lys	Ser	Phe 145	Ile	Asn	Met	Gly	Pro 150	Ile	Lys	Ala						
ATT	ACT	CAG	CAA	TCA	CAA	CAA	TCT	GGT	CAT	TTC	TTA	CAT	TAT	AAA	ATG	1213					
Ile	Thr 155	Gln	Gln	Ser	Gln	Gln	Ser	Gly	His	Phe	Leu 165	His	Tyr	Lys	Met						
ACA	GAA	GGT	ATT	GAA	AGT	ATA	GAG	CGC	TCT	GAT	GGG	ATT	GGC	GAA	ATA	1261					
Thr	Glu	Gly	Ile	Glu	Ser 175	Ile	Glu	Arg	Ser	Asp 180	Gly	Ile	Gly	Glu	Ile 185						
TTC	GAC	CCC	CTA	TAT	GAT	ATT	CTT	TCT	AAG	AAC	GAC	AGA	GCA	ATT	TCA	1309					
Phe	Asp	Pro	Leu 190	Tyr	Asp	Ile	Leu	Ser	Lys 195	Asn	Asp	Arg	Ala	Ile 200	Ser						
AGA	ACT	TTA	AAA	AAA	GAA	GAG	TTA	GAT	CCT	TCC	AGT	AAC	TTC	AAT	AAA	1357					
Arg	Thr	Leu	Lys 205	Lys	Glu	Glu	Leu	Asp 210	Pro	Ser	Ser	Asn	Phe 215	Asn	Lys						
GAC	TCA	GTA	CGA	TTT	ATT	CAT	GAC	GTA	ATT	TTT	GTA	TGT	GGT	CCT	TTG	1405					
Asp	Ser	Val 220	Arg	Phe	Ile	His	Asp 225	Val	Ile	Phe	Val	Cys 230	Gly	Pro	Leu						
CAA	CTT	AAT	GAA	CTC	ATC	GAA	ATA	ATC	ACA	AAA	ATA	TTT	GGC	ACA	GAA	1453					
Gln	Leu 235	Asn	Glu	Leu	Ile	Glu 240	Ile	Ile	Thr	Lys	Ile 245	Phe	Gly	Thr	Glu						
AGC	CAT	TAC	AAA	AAA	AAT	CTT	CTA	AAG	CAC	CTT	GGT	ATT	CTA	ATA	GCT	1501					
Ser 250	His	Tyr	Lys	Lys	Asn 255	Leu	Leu	Lys	His	Leu 260	Gly	Ile	Leu	Ile	Ala 265						
ATT	AGA	ATA	ATA	TCA	TGC	ACA	AAT	GGG	ATT	TAT	TAT	TCT	TTG	TAT	AAA	1549					
Ile	Arg	Ile	Ile	Ser 270	Cys	Thr	Asn	Gly 275	Ile	Tyr	Tyr	Ser	Leu 280	Tyr	Lys						
GAA	TAT	TAT	TTT	AAA	TAT	GAC	TTT	GAC	ATT	GAC	AAC	ATA	TCA	TCA	ATG	1597					
Glu	Tyr	Tyr	Phe 285	Lys	Tyr	Asp	Phe	Asp 290	Ile	Asp	Asn	Ile	Ser 295	Ser	Met						
TTT	AAA	GTT	TTT	TTC	CTC	AAG	AAC	AAG	CCA	GAA	AGG	ATG	AGG	GTA	TAT	1645					
Phe	Lys	Val 300	Phe	Phe	Leu	Lys	Asn 305	Lys	Pro	Glu	Arg	Met 310	Arg	Val	Tyr						
GAG	AAT	ATA	TAGCCTAATT				GATTCTCAGA			CATTGATGAC			TAAGGGATTT				1694				
Glu	Asn	Ile 315																			
GCTTCTGAAG			TAATGCGATC			ACCTGAGCCG			CCAAAAAAT				GGGATATAGC			TAAGAAAAAA		1754			
GGAGGTATGA			GAACAATTTA			TCACCCGTCA			TCAAAAGTTA				AATTAATTCA			ATATTGGTTA		1814			

ATGAATAATG TTTTTTCGAA GCTCCCAATG CATAATGCTG CATATGCATT TGTAAAAAAC 1874  
CGATCAATAA AAAGCAATGC TTTATTACAT GCCGAATCAA AGAATAAGTA TTATGTGAAA 1934  
ATAGATCTCA AAGATTTTTT CCCTTCAATA AAATTTACTG ATTTTGAGTA CGCATTCACT 1994  
CGTTATCGAG ATCGCATTGA ATTTACTACA GAATATGATA AGGAGTTACT ACAACTTATA 2054  
AAAACGATCT GCTTTATATC AGATAGCACT CTCCCTATCG GGTTCCTAC ATCTCCATTA 2114  
ATTGCAAACCT TTGTGGCAAG AGAACTTGAT GAAAAACTGA CGCAAAACCT AAATGCAATT 2174  
GATAAACTTA ATGCCACTTA TACACGATAT GCTGATGATA TTATTGTCTC TACAAATATG 2234  
AAAGGGGCTA GCAAATTAAT TCTGGATTGT TTTAAAAGAA CAATGAAAGA GATTGGTCCA 2294  
GACTTTAAAA TTAACATTAA AAAATTTAAG ATTTGTAGTG CTTCGGGAGG AAGTATAGTA 2354  
GTTACCGGAT TGAAAGTTTG CCACGATTTT CATATTACAT TACATAGATC AATGAAAGAT 2414  
AAAATAAGAT TGCATCTTTC TCTTTTATCA AAGGGCATAT TAAAAGATGA AGATCATAAT 2474  
AAACTTTCTG GTTATATTGC TTATGCAAAA GATATAGACC CTCATTTTTTA TACAAAACCTG 2534  
AACAGAAAAT ATTTTCAAGA AATAAAATGG ATTCAGAATC TCCACAACAA AGTTGAATAA 2594  
ACTTTATATT TTGGATGCAC CCCAATAACT TCATTGATTA AATTGGGAAC AATATAGGCT 2654  
TTTCAGGATG ACCTACACTC TAGAGAATGT GTATACAAAA GTGTATAAGT TATTTTCAAA 2714  
CCTATATAAA ATACAGCAAA ATCAATGCAT TGGCGGCATT TTACCACTCC TGTGATCTTC 2774  
CGCCAAAATG CCTC 2788

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 316 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met	Arg	Ile	Tyr	Ser	Leu	Ile	Asp	Ser	Gln	Thr	Leu	Met	Thr	Lys	Gly
1				5				10						15	
Phe	Ala	Ser	Glu	Val	Met	Arg	Ser	Pro	Glu	Pro	Pro	Lys	Lys	Trp	Asp
			20					25					30		
Ile	Ala	Lys	Lys	Lys	Gly	Gly	Met	Arg	Thr	Ile	Tyr	His	Pro	Ser	Ser
		35					40					45			

Lys Val Lys Leu Ile Gln Tyr Trp Leu Met Asn Asn Val Phe Ser Lys  
 50 55 60  
 Leu Pro Met His Asn Ala Ala Tyr Ala Phe Val Lys Asn Arg Ser Ile  
 65 70 75 80  
 Lys Ser Asn Ala Leu Leu His Ala Glu Ser Lys Asn Lys Tyr Tyr Val  
 85 90 95  
 Lys Ile Asp Leu Lys Asp Phe Phe Pro Ser Ile Lys Phe Thr Asp Phe  
 100 105 110  
 Glu Tyr Ala Phe Thr Arg Tyr Arg Asp Arg Ile Glu Phe Thr Thr Glu  
 115 120 125  
 Tyr Asp Lys Glu Leu Leu Gln Leu Ile Lys Thr Ile Cys Phe Ile Ser  
 130 135 140  
 Asp Ser Thr Leu Pro Ile Gly Phe Pro Thr Ser Pro Leu Ile Ala Asn  
 145 150 155 160  
 Phe Val Ala Arg Glu Leu Asp Glu Lys Leu Thr Gln Lys Leu Asn Ala  
 165 170 175  
 Ile Asp Lys Leu Asn Ala Thr Tyr Thr Arg Tyr Ala Asp Asp Ile Ile  
 180 185 190  
 Val Ser Thr Asn Met Lys Gly Ala Ser Lys Leu Ile Leu Asp Cys Phe  
 195 200 205  
 Lys Arg Thr Met Lys Glu Ile Gly Pro Asp Phe Lys Ile Asn Ile Lys  
 210 215 220  
 Lys Phe Lys Ile Cys Ser Ala Ser Gly Gly Ser Ile Val Val Thr Gly  
 225 230 235 240  
 Leu Lys Val Cys His Asp Phe His Ile Thr Leu His Arg Ser Met Lys  
 245 250 255  
 Asp Lys Ile Arg Leu His Leu Ser Leu Leu Ser Lys Gly Ile Leu Lys  
 260 265 270  
 Asp Glu Asp His Asn Lys Leu Ser Gly Tyr Ile Ala Tyr Ala Lys Asp  
 275 280 285  
 Ile Asp Pro His Phe Tyr Thr Lys Leu Asn Arg Lys Tyr Phe Gln Glu  
 290 295 300  
 Ile Lys Trp Ile Gln Asn Leu His Asn Lys Val Glu  
 305 310 315

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1602 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 548..1507

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CGC  
TGGCATCTAT TAAGAAGGTT AGGAAAGAAA ATAAAGTATC AAAAGATATT GGAAATATAT 60  
TATACGCAGA GCGTTTCTAT TGCCTTGTAT CTATTTACTG GATAGTGTCA ACTACCGCAC 120  
ACTGTGTGAA CTAGCTTTTA AAGCGATAAA GCAAGATGAT GTTTTATCTA AAATTATTGT 180  
TAGATCCGTT GTTTCTCGTC TAATAAATGA ACGAAAAATA CTTCAAATGA CTGATGGTTA 240  
TCAGGTCACT GCTTTGGGGG CTAGCTATGT TAGGAGCGTC TTTGATAGAA AGACACTTGA 300  
CCGATTGCGG CTTGAGATTA TGAATTTTGA AAACCGTAGA AAATCAACAT TTAACATGA 360  
TAAGATTCCG TATGCGCACC CTTAGCGAGA GGTTTATCAT TAAGGTCAAC CTCTGGATGT 420  
TGTTTCGGCA TCCTGCATTG AATCTGAGTT ACTGTCTGTT TTCCTTGTTG GAACGGAGAG 480  
CATCGCCTGA TGCTCTCCGA GCCAACCAGG AAACCCGTTT TTTCTGACGT AAGGGTGCGC 540  
AACTTTC ATG AAA TCC GCT GAA TAT TTG AAC ACT TTT AGA TTG AGA AAT 589  
Met Lys Ser Ala Glu Tyr Leu Asn Thr Phe Arg Leu Arg Asn  
1 5 10  
CTC GGC CTA CCT GTC ATG AAC AAT TTG CAT GAC ATG TCT AAG GCG ACT 637  
Leu Gly Leu Pro Val Met Asn Asn Leu His Asp Met Ser Lys Ala Thr  
15 20 25 30  
CGC ATA TCT GTT GAA ACA CTT CGG TTG TTA ATC TAT ACA GCT GAT TTT 685  
Arg Ile Ser Val Glu Thr Leu Arg Leu Leu Ile Tyr Thr Ala Asp Phe  
35 40 45  
CGC TAT AGG ATC TAC ACT GTA GAA AAG AAA GGC CCA GAG AAG AGA ATG 733  
Arg Tyr Arg Ile Tyr Thr Val Glu Lys Lys Gly Pro Glu Lys Arg Met  
50 55 60  
AGA ACC ATT TAC CAA CCT TCT CGA GAA CTT AAA GCC TTA CAA GGA TGG 781  
Arg Thr Ile Tyr Gln Pro Ser Arg Glu Leu Lys Ala Leu Gln Gly Trp  
65 70 75  
GTT CTA CGT AAC ATT TTA GAT AAA CTG TCG TCA TCT CCT TTT TCT ATT 829  
Val Leu Arg Asn Ile Leu Asp Lys Leu Ser Ser Ser Pro Phe Ser Ile  
80 85 90



GGA	TTT	GAA	AAG	CAC	CAA	TCT	ATT	TTG	AAT	AAT	GCT	ACC	CCG	CAT	ATT	877
Gly	Phe	Glu	Lys	His	Gln	Ser	Ile	Leu	Asn	Asn	Ala	Thr	Pro	His	Ile	
95					100				105						110	
GGG	GCA	AAC	TTT	ATA	CTG	AAT	ATT	GAT	TTG	GAG	GAT	TTT	TTC	CCA	AGT	925
Gly	Ala	Asn	Phe	Ile	Leu	Asn	Ile	Asp	Leu	Glu	Asp	Phe	Phe	Pro	Ser	
				115				120						125		
TTA	ACT	GCT	AAC	AAA	GTT	TTT	GGA	GTG	TTC	CAT	TCT	CTT	GGT	TAT	AAT	973
Leu	Thr	Ala	Asn	Lys	Val	Phe	Gly	Val	Phe	His	Ser	Leu	Gly	Tyr	Asn	
			130				135						140			
CGA	CTA	ATA	TCT	TCA	GTT	TTG	ACA	AAA	ATA	TGT	TGT	TAT	AAA	AAT	CTG	1021
Arg	Leu	Ile	Ser	Ser	Val	Leu	Thr	Lys	Ile	Cys	Cys	Tyr	Lys	Asn	Leu	
		145					150					155				
CTA	CCA	CAA	GGT	GCT	CCA	TCA	TCA	CCT	AAA	TTA	GCT	AAT	CTA	ATA	TGT	1069
Leu	Pro	Gln	Gly	Ala	Pro	Ser	Ser	Pro	Lys	Leu	Ala	Asn	Leu	Ile	Cys	
	160					165					170					
TCT	AAA	CTT	GAT	TAT	CGT	ATT	CAG	GGT	TAT	GCA	GGT	AGT	CGG	GGC	TTG	1117
Ser	Lys	Leu	Asp	Tyr	Arg	Ile	Gln	Gly	Tyr	Ala	Gly	Ser	Arg	Gly	Leu	
175					180					185					190	
ATA	TAT	ACG	AGA	TAT	GCC	GAT	GAT	CTC	ACC	TTA	TCT	GCA	CAG	TCT	ATG	1165
Ile	Tyr	Thr	Arg	Tyr	Ala	Asp	Asp	Leu	Thr	Leu	Ser	Ala	Gln	Ser	Met	
				195				200						205		
AAA	AAG	GTT	GTT	AAA	GCA	CGT	GAT	TTT	TTA	TTT	TCT	ATA	ATC	CCA	AGT	1213
Lys	Lys	Val	Val	Lys	Ala	Arg	Asp	Phe	Leu	Phe	Ser	Ile	Ile	Pro	Ser	
		210					215						220			
GAA	GGA	TTG	GTT	ATT	AAC	TCA	AAA	AAA	ACT	TGT	ATT	AGT	GGG	CCT	CGT	1261
Glu	Gly	Leu	Val	Ile	Asn	Ser	Lys	Lys	Thr	Cys	Ile	Ser	Gly	Pro	Arg	
		225					230					235				
AGT	CAG	AGG	AAA	GTT	ACA	GGT	TTA	GTT	ATT	TCA	CAA	GAG	AAA	GTT	GGG	1309
Ser	Gln	Arg	Lys	Val	Thr	Gly	Leu	Val	Ile	Ser	Gln	Glu	Lys	Val	Gly	
	240					245					250					
ATA	GGT	AGA	GAA	AAA	TAT	AAA	GAA	ATT	AGA	GCA	AAG	ATA	CAT	CAT	ATA	1357
Ile	Gly	Arg	Glu	Lys	Tyr	Lys	Glu	Ile	Arg	Ala	Lys	Ile	His	His	Ile	
255					260					265					270	
TTT	TGC	GGT	AAG	TCT	TCT	GAG	ATA	GAA	CAC	GTT	AGG	GGA	TGG	TTG	TCA	1405
Phe	Cys	Gly	Lys	Ser	Ser	Glu	Ile	Glu	His	Val	Arg	Gly	Trp	Leu	Ser	
				275				280						285		
TTT	ATT	TTA	AGT	GTG	GAT	TCA	AAA	AGC	CAT	AGG	AGA	TTA	ATA	ACT	TAT	1453
Phe	Ile	Leu	Ser	Val	Asp	Ser	Lys	Ser	His	Arg	Arg	Leu	Ile	Thr	Tyr	
			290				295						300			
ATT	AGC	AAA	TTA	GAA	AAA	AAA	TAT	GGA	AAG	AAC	CCT	TTA	AAT	AAA	GCG	1501
Ile	Ser	Lys	Leu	Glu	Lys	Lys	Tyr	Gly	Lys	Asn	Pro	Leu	Asn	Lys	Ala	

SECRET FEB 20 1967

315

1557

1602

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1540 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(A) NAME/KEY: CDS

- (B) LOCATION: 396..1352

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GCGCCTCTTT TGGCCGCCGT GATGTGGAGA GTGGA ATG GAT GCT ACC CGG ACA 413  
Met Asp Ala Thr Arg Thr  
1 5

AAA GAA ATA CAG CGA CTG CAT GCG CTC AGT AAT CAT GCC GGA CGC CAT 509  
Lys Glu Ile Gln Arg Leu His Ala Leu Ser Asn His Ala Gly Arg His  
25 30 35

TAC CGA CGC ATT ATT CTT TCT AAA CGC CAC GGT GGT CAG CGG CTG GTG 557  
Tyr Arg Arg Ile Ile Leu Ser Lys Arg His Gly Gly Gln Arg Leu Val  
40 45 50

TTA GCC CCT GAT TAC TTG CTC AAA ACC GTA CAG CGC AAC ATT CTT AAG 605

Leu 55	Ala	Pro	Asp	Tyr	Leu 60	Leu	Lys	Thr	Val	Gln 65	Arg	Asn	Ile	Leu	Lys 70	
AAC Asn	GTC Val	CTT Leu	TCA Ser	CAA Gln 75	TTT Phe	CCG Pro	CTT Leu	TCC Ser	CCT Pro 80	TTT Phe	GCT Ala	ACA Thr	GCC Ala	TAC Tyr 85	CGA Arg	653
CCA Pro	GGT Gly	TGC Cys	CCA Pro 90	ATC Ile	GTC Val	AGC Ser	AAC Asn	GCG Ala 95	CAG Gln	CCA Pro	CAC His	TGC Cys	CAA Gln 100	CAG Gln	CCG Pro	701
CAG Gln	ATC Ile	CTG Leu 105	AAA Lys	CTC Leu	GAT Asp	ATC Ile	GAA Glu 110	AAC Asn	TTT Phe	TTC Phe	GAT Asp	AGC Ser 115	ATT Ile	AGC Ser	TGG Trp	749
TTA Leu	CAG Gln 120	GTC Val	TGG Trp	CGT Arg	GTG Val	TTT Phe 125	CGC Arg	CAG Gln	GCC Ala	CAG Gln 130	TTG Leu	CCA Pro	CGT Arg	AAT Asn	GTG Val	797
GTA Val 135	ACC Thr	ATG Met	CTG Leu	ACC Thr	TGG Trp 140	ATT Ile	TGT Cys	TGT Cys	TAT Tyr	AAC Asn 145	GAC Asp	GCG Ala	TTA Leu	CCG Pro	CAG Gln 150	845
GGG Gly	GCA Ala	CCA Pro	ACT Thr	TCG Ser 155	CCA Pro	GCC Ala	ATT Ile	TCC Ser	AAT Asn 160	CTT Leu	GTG Val	ATG Met	CGC Arg	CGT Arg 165	TTT Phe	893
GAT Asp	GAA Glu	CGC Arg	ATA Ile 170	GGG Gly	GAA Glu	TGG Trp	TGT Cys	CAG Gln 175	GCT Ala	CGG Arg	GGA Gly	ATT Ile 180	ACC Thr	TAC Tyr	ACC Thr	941
CGC Arg	TAC Tyr	TGC Cys 185	GAT Asp	GAC Asp	ATG Met	ACC Thr	TTT Phe 190	TCA Ser	GGT Gly	CAC His	TTC Phe	AAT Asn 195	GCC Ala	CGC Arg	CAG Gln	989
GTT Val 200	AAA Lys	AAT Asn	AAA Lys	GTG Val	TGC Cys	GGA Gly 205	TTG Leu	TTA Leu	GCG Ala	GAG Glu 210	CTG Leu	GGC Gly	CTG Leu	AGC Ser	CTC Leu	1037
AAT Asn 215	AAA Lys	CGC Arg	AAA Lys	GGC Gly 220	TGC Cys	CTG Leu	ATA Ile	GCT Ala	GCC Ala 225	TGT Cys	AAG Lys	CGC Arg	CAG Gln	CAA Gln 230	GTA Val	1085
ACC Thr	GGG Gly	ATT Ile	GTT Val 235	GTT Val	AAT Asn	CAC His	AAG Lys	CCA Pro	CAG Gln 240	CTT Leu	GCC Ala	CGT Arg	GAA Glu 245	GCG Ala	CGC Arg	1133
CGG Arg	GCG Ala	CTG Leu 250	CGT Arg	CAG Gln	GAG Glu	GTG Val	CAT His	TTG Leu 255	TGC Cys	CAA Gln	AAA Lys	TAT Tyr 260	GGC Gly 260	GTT Val	ATT Ile	1181
TCG Ser	CAT His 265	CTT Leu	AGT Ser	CAT His	CGT Arg	GGT Gly 270	GAA Glu 270	CTT Leu	GAT Asp	CCT Pro	TCT Ser	GGC Gly 275	GAT Asp	CTC Leu	CAC His	1229

GCA CAG GCA ACG GCG TAT CTT TAT GCT TTG CAG GGA AGA ATA AAC TGG	1277
Ala Gln Ala Thr Ala Tyr Leu Tyr Ala Leu Gln Gly Arg Ile Asn Trp	
280 285 290	
TTA TTG CAA ATC AAC CCT GAG GAT GAG GCC TTT CAA CAG GCG AGA GAG	1325
Leu Leu Gln Ile Asn Pro Glu Asp Glu Ala Phe Gln Gln Ala Arg Glu	
295 300 305 310	
AGT GTA AAG CGA ATG CTG GTT GCA TGG TAAGAAAAGC GTCAGGCAGA	1372
Ser Val Lys Arg Met Leu Val Ala Trp	
315	
CGTTTCTGCC TGACCGTTTA GGGGAGAATT ACTGCAACTG CGCGGCAATT AGCGGCCAGC	1432
GGGCGTCAAA ATCATCCGTC GGGCGGTATT TAAACTCGCT GCGGACAAAA CGTGACAGCA	1492
TACCTTCACA GAAGGCCAGG ATCTGGCTTG CCAGCAGGGT TTCATCGG	1540

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Tyr Xaa Asp Asp  
1 4

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Ser Xaa Xaa Xaa  
1 4

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 4 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Xaa Val Thr Gly  
1 4

255050-1030397